IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Patent Application No. 10/649,457

Applicant: Crystal et al.

Filed: August 27, 2003

TC/AU: 1632

Examiner: Marcia Stephens Noble

Docket No.: 216474 (Client Reference No. 3044-01 (US) BK)

Customer No.: 23460

APPELLANTS' APPEAL BRIEF

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In support of the appeal from the final rejection dated September 7, 2007, Appellants now submit their Brief.

Real Party In Interest

The patent application that is the subject of this appeal is assigned to Cornell Research Foundation, Inc.

Related Appeals and Interferences

There are no appeals or interferences that are related to this appeal.

Status of Claims

Claims 1-3, 6-19, 21, and 42-58 are pending. Claims 11 and 12 were withdrawn as being drawn to a non-elected invention in response to a restriction requirement by way of the "Reply to Office Action" dated April 24, 2006. In the "Reply to Office Action" dated August 21, 2006, claims 4, 5, and 22-41 were cancelled, and claims 42-57 were added. In the "Reply to Office Action" dated January 15, 2007, claim 20 was cancelled, claims 2, 3, and 42-57

were withdrawn, and claim 58 was added and withdrawn, as drawn to a non-elected invention. Claims 1, 6-10, 13-19, and 21 are the subject of this appeal.

Status of Amendments

No amendments were made after the final rejection dated September 7, 2007.

Summary of Claimed Subject Matter

The invention defined by appealed claim 1 is directed to a gene transfer vector (see specification at, e.g., page 3, line 10 – page 5, line 19) comprising a nucleic acid sequence which encodes an exotoxin of *Bacillus anthracis* (see specification at, e.g., page 2, lines 26-28 and page 10, lines 18-21) and a nucleic acid sequence which encodes a heterologous sorting signal (see specification at, e.g., page 15, line 19 – page 17, line 3), wherein the nucleic acid sequence encoding the exotoxin comprises SEQ ID NO: 1 (see specification at, e.g., page 17, lines 36-37, and page 26, lines 33-34).

Grounds of Rejection to be Reviewed on Appeal

The only ground of rejection to be reviewed on appeal is whether claims 1, 6-10, 13-19, and 21 comply with the enablement requirement of 35 U.S.C. § 112, first paragraph.

Argument

The final Office Action maintains that the subject matter defined by claims 1, 6-10, 13-19, and 21 allegedly fail to comply with the enablement requirement of 35 U.S.C. § 112, first paragraph.

According to the final Office Action, the present specification enables a serotype 5 adenoviral vector comprising (i) a nucleic acid sequence comprising SEQ ID NO: 1, and (ii) a nucleic acid sequence encoding a cleavable LAMP-1 sorting signal, both of which are operably linked to a CMV IE promoter-enhancer. The present specification, however, allegedly does not enable (i) an adenoviral vector comprising a nucleic acid sequence comprising SEQ ID NO: 1 and any heterologous sorting signal or signal peptide that lacks a promoter sequence, (ii) transduction of antigen presenting cells by the gene transfer vector, and (iii) a pharmaceutical composition comprising the gene transfer vector.

(i) Enablement of a Heterologous Sorting Signal and Signal Peptide

The final Office Action alleges that, because the only operable heterologous sorting signal disclosed in the application is the LAMP-1 sorting signal, the specification only enables the use of the LAMP-1 sequence. The final Office Action also contends that the scope of the appealed claims is so broad that one of ordinary skill in the art would not know how to direct the extoxin to antigen presenting cells in order to elicit an immune response.

The enablement standard does not require that Appellants disclose every operable embodiment of a particular invention. Instead, the test of enablement requires that an application contain sufficient information regarding the subject matter of the claims so as to enable one skilled in the art to make and use the claimed invention without undue experimentation. Moreover, a specification need not disclose, and preferably omits, what is well-known to those skilled in the art. *In re Buchner*, 929 F.2d 660, 661, 18 U.S.P.Q. 2d 1331, 1332 (Fed. Cir. 1991); *Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384, 231 U.S.P.Q. 81, 94 (Fed. Cir. 1986), *cert. denied*, 480 U.S. 947 (1987); *Lindemann Maschinenfabrik GMBH v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1463, 221 U.S.P.Q. 481, 489 (Fed. Cir. 1984).

The structures and sequences of numerous sorting signals and signal peptides, particularly sorting signals that direct proteins to lysosomes, were well known in the art well before the present application was filed. For example, Thomas et al., *J. Cell. Sci.*, *116*(Pt 11): 2213-22 (2003), discloses the sequence of the sorting signal of the connexin43 protein which directs it to lysosomes. Klionsky et al., *J. Biol. Chem.*, *265*(10): 5349-52 1990, discloses the sequence of sorting signals for the vacuolar membrane protein repressible alkaline phosphatase. Hogue et al., *Biochem. J.*, *365*(Pt 3): 721-730 (2002), discloses the sequence of the sorting signal of lysosome-associated protein transmembrane 4 alpha (LAPTM4 alpha). Sequences of numerous signal peptides are disclosed in, for example, Alberts et al. (eds.), *Molecular Biology of the Cell,* 3rd Edition, Garland Publishing Inc, New York (1994), p. 558, Lehninger et al. (eds.), *Principles of Biochemistry*, 2nd Edition, Worth Publishers, New York (1993), p. 929, and Nothwehr et al., *Bioessays, 12*: 479-484 (1990).

Thus, using only routine molecular biology techniques, one of ordinary skill in the art would have been able to make and use the claimed invention using any suitable sorting signal

in view of the disclosure of the present specification coupled with the knowledge in the art at the time the application was filed.

(ii) Enablement of Promoter Sequences

According to the final Office Action, because claim 1 does not require operable linkage to a promoter, one of ordinary skill in the art would not know how to use a nucleic acid sequence comprising SEQ ID NO: 1 without conducting undue experimentation. The Office Action also contends that, because the claims do not require a specific promoter, the gene transfer vector defined by the appealed claims would be nonfunctional.

Again, a claim need not recite every operable embodiment in order to satisfy the enablement requirement. The fact that claim 1 does not recite that the nucleic acid sequence is operably linked to a promoter does not preclude enablement of claim 1. The level of skill in the art is such that an ordinarily skilled artisan would understand that, in order to express a particular gene in a cell, the gene must be linked to a promoter. In addition, the specification discloses numerous examples of promoters suitable for use in the invention (see, e.g., paragraph 0037). In addition, Appellants agree with the Examiner's contention that one of ordinary skill in the art would be well aware that a gene transfer vector lacking a promoter to control expression of a transgene would be nonfunctional, which is precisely why one of ordinary skill in the art would understand that the gene transfer vector defined by the appealed claims must contain a promoter. Thus, one of ordinary skill in the art would have been more than adequately equipped to make and use the claimed gene transfer vector given the high level of skill in the art, and the disclosure of the subject application.

(iii) Enablement of Transducing Antigen Presenting Cells

With respect to appealed claim 19, the final Office contends that the specification does not teach any means for targeting the claimed gene transfer vector to antigen presenting cells *in vivo*. However, the specification discloses that adenovirus infection of dendritic cells (DC) is mediated by the interaction of the RGD domain of the penton base protein and integrin molecules expressed at the DC cell surface (see paragraph 0048). In addition, the specification discloses that dendritic cells can be targeted by the adenoviral vector by modifying the fiber protein to contain an RGD domain (see paragraph 0048). The specification also discloses that an adenovirus containing deletions of the fiber CAR domain and penton RGD domain recognizes and infects antigen presenting cells in the liver more

efficiently than an adenovirus comprising wild-type coat proteins (see paragraph 0049). Thus, the present specification enables the subject matter of claim 19.

(iv) Enablement of a Pharmaceutical Composition

The final Office Action alleges that claim 21 is not enabled by the present specification because the specification does not demonstrate that the claimed gene transfer vector produces an immune response *in vivo*.

Whether or not the claimed gene transfer vector produces an immune response *in vivo* is of no consequence to the enablement of a pharmaceutical composition comprising the gene transfer vector and a pharmaceutically acceptable carrier. Appellants note that the specification discloses how to prepare a composition comprising the claimed gene transfer vector and a pharmaceutically acceptable carrier (see, e.g., specification at paragraphs 0054 to 0063). In addition, the specification discloses methods for administering the composition to a mammal to induce an immune response (see, e.g., Examples 2-4).

In view of the foregoing, the invention defined by 1, 6-10, 13-19, and 21 is described in the specification such that one of ordinary skill in the art would be able to make and use the invention. Accordingly, the Section 112, first paragraph, rejection should be reversed.

Respectfully submitted,

Melissa E. Kolom, Reg. No. 51,860
LEYDIG, VOIT & MAYER, LTD.
Two Prudential Plaza
180 North Stetson Ave., Suite 4900
Chicago, Illinois 60601-6731
(312) 616-5600 (telephone)
(312) 616-5700 (facsimile)

Date: January 8, 2008

Claims Appendix

1. (Previously Presented) A gene transfer vector comprising a nucleic acid sequence which encodes an exotoxin of *Bacillus anthracis* and a nucleic acid sequence which encodes a heterologous sorting signal, wherein the nucleic acid sequence encoding the exotoxin comprises SEQ ID NO: 1.

2. (Withdrawn) The gene transfer vector of claim 1, wherein the nucleic acid sequence further comprises a sequence that encodes at least an immunogenic portion of one or more exotoxins selected from the group consisting of edema factor, and lethal factor.

3. (Withdrawn) The gene transfer vector of claim 2, wherein the nucleic acid sequence comprises a sequence selected from the group consisting of SEQ ID NO: 2 and SEQ ID NO: 3.

4.-5. (Canceled)

6. (Original) The gene transfer vector of claim 1, wherein the heterologous sorting signal directs the exotoxin to a subcellular sorting pathway.

7. (Previously Presented) The gene transfer vector of claim 6, wherein the subcellular sorting pathway is a lysosome pathway.

8. (Original) The gene transfer vector of claim 1, wherein the heterologous sorting signal is a lysosomal-associated membrane protein-1 sorting signal.

9. (Original) The gene transfer vector of claim 1, wherein the nucleic acid sequence further encodes a heterologous signal peptide.

10. (Original) The gene transfer vector of claim 9, wherein the heterologous signal peptide is a lysosomal-associated membrane protein-1 signal peptide.

11. (Withdrawn) The gene transfer vector of claim 1, which is a non-viral vector.

12. (Withdrawn) The gene transfer vector of claim 11, wherein the non-viral vector is a plasmid formulated with a lipid or a polymer.

13. (Original) The gene transfer vector of claim 1, which is a viral vector.

14. (Original) The gene transfer vector of claim 13, wherein the viral vector is an adenoviral vector.

15. (Original) The gene transfer vector of claim 14, wherein the adenoviral vector is replication-deficient.

16. (Original) The gene transfer vector of claim 15, wherein the adenoviral vector is a human adenoviral vector.

17. (Original) The gene transfer vector of claim 15, wherein the adenoviral vector is a non-human primate adenoviral vector.

18. (Original) The gene transfer vector of claim 17, wherein the adenoviral vector is a chimpanzee adenoviral vector.

19. (Original) The gene transfer vector of claim 1, wherein the gene transfer vector transduces antigen presenting cells.

20. (Cancelled)

21. (Previously Presented) A composition comprising the gene transfer vector of claim 1 and a pharmaceutically acceptable carrier.

22.-41. (Canceled)

42. (Withdrawn) A method of producing an immune response against *Bacillus anthracis* in a host, which method comprises administering to the host the gene transfer vector of claim 1, and wherein the nucleic acid sequence is expressed to produce the

immunogenic portion of the one or more exotoxins in the host, thereby producing an immune response against *Bacillus anthracis*.

43. (Withdrawn) The method of claim 42, wherein the heterologous sorting signal directs the exotoxin to a subcellular sorting pathway.

44. (Withdrawn) The method of claim 43, wherein the subcellular sorting pathway is selected from the group consisting of an extracellular pathway, a cytoplasmic pathway, a cell membrane pathway, a lysosome pathway, an endoplasmic reticulum pathway, and a degradative pathway.

45. (Withdrawn) The method of claim 42, wherein the heterologous sorting signal is a lysosomal-associated membrane protein-1 sorting signal.

46. (Withdrawn) The method of claim 42, wherein the nucleic acid sequence further encodes a heterologous signal peptide.

47. (Withdrawn) The method of claim 46, wherein the heterologous signal peptide is a lysosomal-associated membrane protein-1 signal peptide.

48. (Withdrawn) The method of claim 42, wherein the gene transfer vector is a non-viral vector.

49. (Withdrawn) The method of claim 48, wherein the non-viral vector is a plasmid formulated with a lipid or a polymer.

50. (Withdrawn) The method of claim 42, wherein the gene transfer vector is a viral vector.

51. (Withdrawn) The method of claim 50, wherein the viral vector is an adenoviral vector.

52. (Withdrawn) The method of claim 51, wherein the adenoviral vector is replication-deficient.

53. (Withdrawn) The method of claim 51, wherein the adenoviral vector is a human adenoviral vector.

54. (Withdrawn) The method of claim 51, wherein the adenoviral vector is a non-human primate adenoviral vector.

55. (Withdrawn) The method of claim 54, wherein the adenoviral vector is a chimpanzee adenoviral vector.

56. (Withdrawn) The method of claim 42, wherein the gene transfer vector is administered to antigen presenting cells of the host.

57. (Withdrawn) The method of claim 56, wherein the antigen presenting cells are dendritic cells.

58. (Withdrawn) The gene transfer vector of claim 6, wherein the subcellular sorting pathway is selected from the group consisting of an extracellular pathway, a cytoplasmic pathway, a cell membrane pathway, an endoplasmic reticulum pathway, and a degradative pathway.

Evidence Appendix

1. Thomas et al., *J Cell Sci.*, *116*(Pt 11): 2213-22 (2003); entered into the record by the Examiner by way of the final Office Action dated September 7, 2007
2. Klionsky et al., *J Biol Chem.*, *265*(10): 5349-52 (1990); entered into the record by the Examiner by way of the final Office Action dated September 7, 2007
3. Hogue et al., *Biochem. J.*, *365*(Pt 3): 721-730 (2002); entered into the record by the Examiner by way of the final Office Action dated September 7, 2007
4. Alberts et al. (eds.), *Molecular Biology of the Cell,* 3rd Edition, Garland Publishing Inc, New York (1994), p.558; entered into the record by the Examiner by way of the final Office Action dated September 7, 2007
5. Lehninger et al. (eds.), *Principles of Biochemistry*, 2nd Edition, Worth Publishers, New York (1993), p.929; entered into the record by the Examiner by way of the final Office Action dated September 7, 2007
6. Nothwehr et al., *Bioessays, 12*: 479-484 (1990); entered into the record by the Examiner by way of the final Office Action dated September 7, 2007

Related Proceedings Appendix

Not applicable

A tyrosine-based sorting signal is involved in connexin43 stability and gap junction turnover

Marc A. Thomas[1,2,3], Nathalie Zosso[1], Isabelle Scerri[2], Nicolas Demaurex[3], Marc Chanson[2,*,‡] and Olivier Staub[1,*,‡]

[1]Institute of Pharmacology and Toxicology, University of Lausanne, 1005 Lausanne, Switzerland
[2]Department of Pediatrics, University Hospitals, 1211 Geneva, Switzerland
[3]Department of Physiology, University of Geneva, 1211 Geneva, Switzerland
*These authors contributed equally to this study
‡Authors for correspondence (e-mails: olivier.staub@ipharm.unil.ch; marc.chanson@houge.ch)

Accepted 17 February 2003
Journal of Cell Science 116, 2213-2222 © 2003 The Company of Biologists Ltd
doi:10.1242/jcs.00440

Summary

The gap junction protein connexin43 is known to have a rapid turnover, involving degradation by both the proteasomal and lysosomal systems, but the structural features of connexin43 that govern these actions are not known. The connexin43 C-terminal sequence contains a proline-rich region corresponding to the consensus of a protein-protein interaction PY-motif (xPPxY), and an overlapping putative tyrosine-based sorting signal (Yxxϕ; ϕ=hydrophobic), known to play a role in the intracellular trafficking of many membrane proteins. As both motifs may control turnover of connexin43, we used a combination of metabolic radiolabelling, immunoprecipitation and functional assays to determine the possible role of these motifs in controlling degradation of human connexin43 expressed in SKHep1 cells. Mutation V289D in the tyrosine-based sorting motif increased the steady-state pool of connexin43 by approximately 3.5-fold, while mutation P283L in the PY-motif produced a comparatively modest augmentation (1.7-fold). No additive effect was observed when the overlapping tyrosine was mutated. In pulse-chase experiments, the Y286A substitution increased the half-life of connexin43 from 2 to 6 hours, indicating that the increased steady-state levels reflected reduced protein degradation. Moreover, expression at the junctional membrane, as well as gap junction-mediated intercellular communication (GJC), were nearly abolished by lysosomal inhibitors and Brefeldin A in cells expressing wild-type connexin43, but were unaffected in the tyrosine mutant. These results provide strong evidence that the tyrosine-based motif of human connexin43 is a prime determinant controlling connexin43 stability, and consequently GJC, by targeting connexin43 for degradation in the endocytic/lysosomal compartment.

Key words: Connexins; Gap junctional communication; Trafficking; Degradation; Endocytosis

Introduction

Gap junctions are complexes of transmembrane proteins, called connexins, that directly link the cytoplasm of adjacent cells. These intercellular channels act as molecular sieves, allowing the passage of low-molecular-weight substances such as amino acids, ions and second messengers (Goodenough et al., 1996; Kumar and Gilula, 1996; Spray, 1996). Gap junctions play important roles in a variety of cellular processes such as embryonic development, synchronous contraction of cardiac and smooth muscle cells, regulation of exocytosis and control of cell growth. Mutations in connexins or defective production of gap junctions are associated with deafness, Charcot-Marie-Tooth X-linked neuropathy, malignancy, cataractogenesis and skin diseases (Rabionet et al., 2002; Simon and Goodenough, 1998; Yamasaki and Naus, 1996).

Unlike most membrane proteins, gap junctions are dynamic structures with half-lives ranging from 1.5 to 5 hours (Darrow et al., 1995; Laird et al., 1991; Musil et al., 1990a). The life cycle of connexins involves the noncovalent oligomerisation of subunits into connexons, the translocation of assembled connexons to the cell surface, intercellular pairing of connexons and channel clustering into paracrystallin arrays referred to as gap junction plaques (reviewed in (Kumar and Gilula, 1996). The retrieval of gap junction plaques from the cell surface has been proposed to entail the endocytosis of partial or complete junctional plaques as a double membrane annular junction that is subsequently degraded or possibly re-utilised (Gaietta et al., 2002; Jordan et al., 2001; Larsen et al., 1979; Naus et al., 1993). Hence, processes involving turnover and degradation, as well as remodelling, may provide important mechanisms to regulate intercellular communication under normal or pathological conditions (Beardslee et al., 1998; Laird, 1996; Luke and Saffitz, 1991; Traub et al., 1983).

Connexin43 (Cx43), the most studied gap junction protein so far, undergoes several types of post-translational modifications, including phosphorylation and ubiquitination. Numerous studies have established that the conversion of unphosphorylated Cx43 to slower migrating species on SDS-polyacrylamide gel is caused by phosphorylation of Cx43, an event that facilitates gap junction channel formation and gating (Kwak et al., 1995; Laird et al., 1991; Lampe, 1994; Moreno et al., 1994; Musil et al., 1990b). The phosphorylation state of Cx43 has also been proposed to control Cx43 degradation in rat mammary tumour cells and in intact rat heart (Beardslee et

al., 1998; Laird et al., 1995). The degradation of Cx43 has been shown to occur by the lysosomal (Laing et al., 1997; Larsen and Hai, 1978; Musil et al., 2000; Naus et al., 1993; Vaughan and Lasater, 1990) and the ubiquitin-proteasomal (Laing et al., 1997; Musil et al., 2000; Laing and Beyer, 1995; Rutz and Hulser, 2001) pathways, the relative contribution of which appears to be largely cell-type specific. It is likely that some of the proteasomal degradation occurs at the level of the endoplasmic reticulum (ER), as a quality control step, to remove poorly folded or oligomerised connexin polypeptides (Musil et al., 2000; VanSlyke et al., 2000).

Despite the large body of knowledge on the rapid turnover and degradation of Cx43, and the proteolytic systems involved, virtually nothing is known about the signals and motifs that control sorting to the lysosome or promote degradation by the ubiquitin-proteasome system. Cx43 does not have an amino terminus (basic or hydrophobic, bulky amino acids) that would be recognised by the N-end rule (Varshavsky, 1992). Many rapidly degrading proteins contain PEST sequences (rich in proline, glutamic acid, serine and threonine), which have been suggested to be signals for rapid degradation (Rogers et al., 1986). Indeed, low consensus PEST sequences were described for Cx43 (Darrow et al., 1995; Laird et al., 1991); however, their role in Cx43 turnover has not been shown experimentally. Interestingly, Cx43 contains a proline-rich motif in its C-terminus, which conforms to the consensus of a PY motif (xPPxY, P=proline, Y=tyrosine, x=amino acid) (Fig. 1A). Such PY motifs have been shown to act as ligands for WW domain-containing proteins (Chen and Sudol, 1995). More importantly, several different ion channels interact with members of the Nedd4/Nedd4-like family of ubiquitin-protein ligases, via PY motif/WW domain interactions, leading to their ubiquitination-dependent downregulation at the plasma membrane (Abriel et al., 2000; Abriel et al., 1999; Schwake et al., 2001; Staub et al., 1996). Hence, it is possible that this motif directs the ubiquitination of Cx43 and plays a role in its targeting for endocytosis and destruction. Overlapping the PY-motif, however, is a tyrosine-based sorting signal conforming to the consensus Yxxϕ (where Y is a tyrosine, x is any amino acid and ϕ is an amino acid with a bulky hydrophobic side chain). Signals of this nature, contained in the cytosolic domains of many plasma membrane proteins, are also known to mediate internalisation and lysosomal targeting for degradation. In addition, some Yxxϕ motifs can direct traffic within the endosomal and late secretory pathways (Bonifacino and Dell'Angelica, 1999; Kirchhausen et al., 1997; Owen and Evans, 1998). The specificity of these processes is believed to be achieved through the interaction of these signals with alternative adapter complex molecules that associate with different protein-sorting machineries (Bonifacino and Dell'Angelica, 1999).

To better understand the molecular mechanisms that control connexin retrieval from the plasma membrane and degradation, we studied the possible contribution of the PY-motif and its overlapping tyrosine-based motif on Cx43 degradation. To this end, we studied the effects of amino acid substitutions within this region on protein stability and sensitivity to proteasome and lysosome inhibitors, as well as on functional expression. We report here that the tyrosine-based sorting signal is a primary element in this region controlling Cx43 turnover.

Materials and Methods

Constructs and plasmids

The full-length human Cx43 cDNA, supplied as a gift from D. C. Spray (Dept of Neuroscience, Albert Einstein College of Medicine, NY), was used to develop Cx43 mutant constructs. Site-directed mutagenesis was performed using a PCR-based technique (Nelson and Long, 1989) to substitute tyrosine 286 with alanine (Cx43-Y286A) or phenylalanine (Cx43-Y286F), proline 283 with leucine (Cx43-P283L), glycine 285 with alanine (Cx43-G285A), and valine 289 with aspartate (Cx43-V289D) either alone or in combination with the proline 283 to leucine mutation (Cx43-P283L/V289D), which were used for transfection studies. The cDNA clones were cloned into the pRC/CMV plasmid (Invitrogen, Carlsbad, CA). Correct mutagenesis was assessed by DNA sequencing (Microsynth, Balgach, Switzerland).

Cell culture and transfection

SKHep1 cells were purchased from the American Type Culture Collection (Rockville, MD) and maintained in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10% FCS and 0.6% penicillin/streptomycin (Invitrogen) in an atmosphere of humidified air/5% CO_2 at 37°C. For the development of the stable cell lines Cx43-WT (wildtype) and Cx43-Y286A, plasmids were transfected into SKHep1 cells using the Effectene transfection reagent (Qiagen, Hilden, Germany), and transformants were selected for neomycin resistance using 400 μg/ml of G418. Likewise, transient transfections were carried out using the Effectene transfection kit, and plasmids were allowed to express Cx43 for 48 hours before SDS-PAGE and western blot analysis.

Western blot analysis

SKHep1 cells, stably or transiently transfected with Cx43 constructs, were washed twice in cold PBS, lysed in radioimmunoprecipitation (RIPA) buffer pH 8.0 (20 mM Tris-HCl, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.5% deoxycholate, 1 mM PMSF, 10 μg/ml leupeptin, 10 μg/ml pepstatin A, 10 μg/ml aprotinin) and the cells harvested on ice by scraping. After centrifugation at 4°C for 5 minutes at 20,000 ***g***, the supernatants were recovered and samples denatured by heating at 95°C for 5 minutes in protein sample buffer (295 mM sucrose, 2% SDS, 2.5 mM EDTA, 62.5 mM Tris-Cl pH 8.8, 0.05% bromophenol blue, 26 mM dithiothreitol (DTT)). Dephosphorylation experiments were performed by treating the cellular lysates with 50 units of calf intestinal phosphatase (New England Biolabs, Beverly, MA) for 3 hours at 37°C before denaturing in protein sample buffer. Cellular lysates were then electrophoresed on a 10% polyacrylamide gel (SDS-PAGE), before being transferred onto a nitrocellulose membrane. Membranes were then probed with a polyclonal antibody directed against the C-terminus of Cx43 (Zymed, San Francisco, CA) or β-Actin (Sigma). For experiments with protease inhibitors, the stable cell lines Cx43-WT and Cx43-Y286A were treated for 3 hours with either lactacystin (10 μM), leupeptin (10 μM) or NH_4Cl (10 mM) before cell harvesting and western blotting. Quantitation of recognised levels was performed on fluorograms, using a molecular imager FX (Biorad, Hercules, CA), and the results were normalised to the controls and expressed as mean±s.e.m. Statistical analyses were performed using the unpaired two-tailed Student's *t* test.

Pulse-chase analysis

Cx43-WT and Cx43-Y286A cells were grown to 80% confluency and then starved in depletion medium (DMEM without methionine) for 30 minutes at 37°C. Cells were then labelled for 60 minutes in depletion medium containing 0.1 mCi/ml [^{35}S]-methionine. After labelling, cells were placed on ice and washed three times in ice-cold wash medium (DMEM, 10% FCS, 0.6% penicillin/streptomycin, 10

mM methionine) before being chased, for various periods of time, in pre-warmed wash medium (alone or supplemented with either 10 μM lactacystin or 10 mM NH_4Cl). At the end of the chase, cells were transferred onto ice, washed three times with ice-cold PBS, then lysed in RIPA buffer. Cells were harvested by scraping, centrifuged for 5 minutes at 20,000 *g* (4°C) and the supernatants immunoprecipitated overnight (4°C with rotation) using an anti-Cx43 antibody (Zymed) together with protein A sepharose beads. After immunoprecipitation, the beads were washed four times in RIPA buffer and then the immunoprecipitated proteins eluted with protein sample buffer and boiling for 5 minutes at 95°C. Radiolabelled proteins were then visualised by SDS-PAGE and autoradiography. Quantitation of [^{35}S]-labelled Cx43 was performed on autoradiographs using a molecular imager FX (Biorad). The results were normalised to the control (t=0) and expressed as mean±s.e.m. Mono-exponential curves for each independent experiment were fitted through the values measured at each time point according to the formula $y=100*\exp(-kt)$, using Kaleidagraph v. 3.52 (Synergy Software, Reading, PA). The decay rate constants (k), which are representative of the rates of protein degradation, were determined for each curve and the mean values calculated. For statistical analyses, two-way ANOVA was performed using Prism (GraphPad, San Diego, CA).

Immunofluorescence microscopy

For immunofluorescent labelling, cell lines were cultured on glass coverslips and incubated for 3 hours in new medium alone (control) or supplemented with 10 μM lactacystin, 10 mM NH_4Cl or 2 μg/ml Brefeldin A before fixation for 2-3 minutes in methanol at –20°C. The coverslips were then rinsed in PBS and incubated successively with 0.2% Triton X-100 for 60 minutes, 0.5 M NH_4Cl for 15 minutes and PBS supplemented with 2% bovine serum albumin for an additional 30 minutes. Cells were then rinsed and incubated overnight with polyclonal antibodies (diluted 1:30) against Cx43 (Alpha Diagnostics, San Antonio, TX). After washing in PBS, the coverslips were incubated with secondary antibodies, conjugated to FITC for 3 hours and then examined using fluorescent microscopy. Images were acquired with a high-sensitivity CCD Visicam (Visitron systems GmbH, Germany) camera connected to a personal computer. Images were captured using the software Metafluor 4.01 (Universal Imaging, West Chester, PA) and processed using Adobe Photoshop 5.5 (Adobe Systems Inc.).

Cell-coupling measurements

Dye coupling studies were performed on subconfluent monolayers of cells incubated in a solution (external solution) containing (in mM): 136 NaCl, 4 KCl, 1 $CaCl_2$, 1 $MgCl_2$, and 2.5 glucose, and was buffered to pH 7.4 with 10 mM HEPES-NaOH. Single cells were impaled with microelectrodes backfilled with a 4% lucifer yellow solution prepared in 150 mM LiCl (buffered to pH 7.2 with 10 mM Hepes). The fluorescent tracer was allowed to fill the cells by simple diffusion for 3 minutes. After the injection period, the electrode was removed and the number of fluorescent cells was counted. Cells were visualised using epifluorescent illumination provided by a 100 W mercury lamp and the appropriate set of filters. The results were expressed as mean±s.e.m. To examine the effects of proteasome/lysosome inhibitors on dye coupling, subconfluent monolayers of cells were incubated for 3 hours in the external solution supplemented with either 10 μM lactacystin, 10 mM NH_4Cl or 2 μg/ml Brefeldin A.

For electrical coupling studies, the dual whole-cell patch-clamp approach was applied to pairs of cells incubated in the external solution. Both cells of a pair were voltage clamped at a common holding potential of 0 mV. To measure gap junctional currents (I_j), transjunctional potential differences (V_j) were elicited by changing the holding potential of one member of a cell pair. I_j was defined as the current recorded in the cell kept at a 0 mV. Junctional conductance (g_j) was then calculated by $g_j=I_j/V_j$ and the results displayed as a scattered plot displaying the individual g_j values, including the mean±s.e.m. Series resistance was not compensated for and was less than 2% of the combined junctional and cell input resistance. Patch electrodes were filled with a pCa 7 solution containing (in mM): 138 KCl, 1 NaCl, 2.9 $CaCl_2$, 5.5 EGTA, 2 $MgCl_2$, and buffered to pH 7.2 with 10 mM HEPES-KOH. Statistical analyses were performed using the two-tailed Student's *t* test for unpaired data.

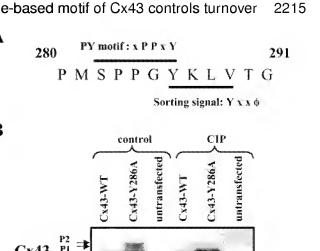


Fig. 1. Mutation of tyrosine 286 to alanine affects Cx43 steady-state levels. (A) The amino acid sequence of Cx43 spanning residues 280-291. The putative PY-motif and tyrosine-based sorting signal are indicated and their consensus sequences displayed. (B) Transient expression of human Cx43 constructs in SKHep1. Cells were transfected with cDNA plasmids encoding either Cx43-WT, Cx43 with a substitution of tyrosine to alanine at position 286 (Cx43-Y286A) or neither (untransfected). Following expression (48 hours), western blot analysis was performed on cellular lysates using either an anti-Cx43 antibody (top panel) or a β-actin antibody (bottom panel) to control protein loading. Cx43 was detected in its nonphosphorylated (NP) and phosphorylated (P1, P2) forms (indicated by arrows) and was determined by treating cellular lysates with CIP before western blot analysis.

Results

Tyrosine 286 is involved in the turnover of connexin43

Inspection of the C-terminus of Cx43 reveals the existence of a PY-motif (xPPxY), known to serve as a ligand for WW domain-containing proteins such as the members of the Nedd4/Nedd4-like family of ubiquitin-protein ligases (Rotin et al., 2000). Adjacent to this PY-motif lies a consensus tyrosine-based sorting signal (Yxxϕ), which may contribute to protein degradation by directing traffic to the endosomal/lysosomal compartment (Fig. 1A). To assess whether gap junction stability is mediated by these motifs, we substituted the tyrosine 286 (Y286), essential to the function of both elements (Chen and Sudol, 1995; Ohno et al., 1995), with an alanine (Y286A), and investigated the effect of this substitution on Cx43 steady-state levels. cDNAs encoding the human Cx43 wild-type protein (Cx43-WT) and its mutant counterpart

(Cx43-Y286A) were transiently transfected into the hepatoma cell line SKHep1, which does not express Cx43 endogenously. Cx43 expression, driven by the cytomegalo virus (CMV) promoter, was then assessed by SDS-PAGE and western blot analysis using an anti-Cx43 antibody. As shown in Fig. 1B, Cx43 could be detected in three different forms, which were not present in untransfected cells. The lower, and most abundant band, corresponds to the non-phosphorylated form of Cx43 (NP), whereas the two slower migrating species, which were most apparent in the Cx43-Y286A mutant, represent the phosphorylated forms, P1 and P2, as confirmed by treatment with calf intestinal phosphatase (Fig. 1B; CIP treatment). Longer exposure also shows the presence of P forms in Cx43-WT-expressing cells (data not shown). Interestingly, the abundance of Cx43 detected in cells transfected with tyrosine-mutated Cx43 was consistently higher than in those transfected with the wildtype. Quantitation of the total pool of Cx43 revealed the increase in Cx43 protein levels to be approximately 3.5-fold (Fig. 2B). A similar result was also observed when the tyrosine was substituted with a phenylalanine (data not shown). This suggested that one of the two overlapping motifs, or both, may control Cx43 steady-state levels. To determine whether the tyrosine residue was important in the context of the PY-motif or the adjacent Y-based motif, we generated a series of substitution mutations around Y286. These included substitution of proline 283 with leucine (Cx43-P283L), glycine 285 with alanine (Cx43-G285A), and valine 289 with aspartate (Cx43-V289D). These constructs were then transiently transfected into SKHep1 cells and Cx43 expression assayed by SDS-PAGE and western blot analysis. As shown in Fig. 2A, the mutation of valine 289, an essential hydrophobic residue within the tyrosine-based sorting signal, mimicked the Y286A mutation. Indeed, quantitative analysis confirmed the increased steady-state expression of Cx43 to be at levels comparable to Cx43-Y286A (Fig. 2B). But mutating the proline residue, a crucial amino acid in the PY-motif, had a less pronounced effect that resulted in a 1.7-fold increase in Cx43 levels. The mutation of glycine 285, which is not required for the sorting signal nor the PY-motif, had no observable effect. We note that mutation of the tyrosine, being part of both the PY-motif and the Y-based motif, showed no additive effect when compared with the valine mutant. Similarly, no additive effect was observed in a P283L/V289D double mutant (data not shown). These results therefore suggest that the tyrosine-based motif is the major element in this region controlling Cx43 protein levels.

The increased expression level of Cx43-Y286A could be due to differences either in the rate of biosynthesis or in the degradation of mutant Cx43, as compared with Cx43-WT. To discriminate between these two possibilities, we developed stable SKHep1 cell lines expressing either Cx43-wildtype or Cx43-Y286A, which was representative of the tyrosine sorting mutations, and examined the turnover of these proteins by pulse-chase analysis. Cells were metabolically labelled for 60 minutes in the presence of [^{35}S]-methionine and then chased for 0, 2, 4 and 6 hours with an excess of unlabelled methionine. Cx43 was then immunoprecipitated and the amount of radiolabelled protein analysed by SDS-PAGE and fluorography. The results, depicted in Fig. 3A, revealed that the degradation of Cx43, harbouring the Y286A mutation, was considerably slower than that of the wild-type Cx43. The two

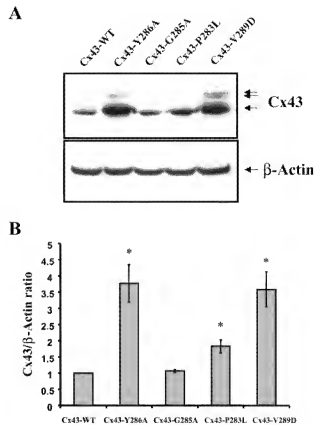


Fig. 2. The tyrosine-based sorting signal, adjacent to the PY-motif, is the major element affecting Cx43 levels. (A) Transient expression of Cx43 mutant constructs in SKHep1 cells. Mutant Cx43 plasmids were developed by exchanging proline with leucine at position 283 (Cx43-P283L), glycine with alanine at position 285 (Cx43-G285A), and valine with aspartate at 289 (Cx43-V289D). Cells were transfected with these plasmids, as well as Cx43-WT and Cx43-Y286A, cellular lysates prepared (48 hours after transfection) and western blot analysis performed using an anti-Cx43 antibody (top panel) or a β-actin antibody (bottom panel) to control protein loading. (B) Quantitation of transfected Cx43 transiently expressed in SKHep1. Experiments were performed as in A, and the levels of total Cx43 and β-Actin detected by western blot fluorography quantified on a molecular imager FX. The Cx43:actin ratio value was determined for each construct and displayed as mean±s.e.m. (n=4 separate experiments). Asterisks indicate differences at P<0.01 vs control as determined by the Student's t-test.

proteins had comparable translation rates as judged by the similar amounts of [^{35}S]-methionine incorporation at time 0. Using this approach, however, Cx43 was predominantly detected as a single band which, according to its predicted molecular weight, corresponds to the NP-form. This suggests that either this assay is not sensitive enough to detect the less-abundant P forms or, alternatively, that immunoprecipitation could not isolate the phosphorylated species. To ensure that the disappearance of the NP form in the pulse-chase experiments represented degradation and not maturation into the phosphorylated forms, we also performed similar experiments in the presence of calf intestinal phosphatases so as to analyse the entire pool of Cx43, and found identical results (data not shown). The [^{35}S]-labelled Cx43 species were quantified in

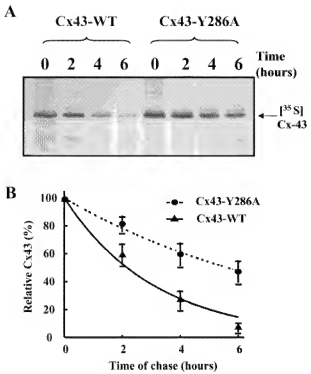


Fig. 3. Tyrosine 286 is involved in Cx43 stability. (A) Pulse-chase analysis was performed on stably transfected SKHep1 cells expressing either Cx43-WT or Cx43-Y286A, by pulsing for 60 minutes with [^{35}S]-methionine and then chasing for 0, 2, 4 or 6 hour periods. Cells were then lysed, Cx43 immunoprecipitated overnight, and SDS-PAGE and fluorography performed. (B) Pulse-chase experiments were performed as in A, and the level of immunoprecipitated [^{35}S]-labelled Cx43 was quantified from fluorographs using a molecular imager FX. A plot of the mean±s.e.m. percentage of pulse-labelled Cx43 remaining after 0, 2, 4 and 6 hours of chase from four independent experiments fitted using a mono-exponential decay function is shown. ANOVA statistical analysis indicates that the difference in the degradation curves between Cx43-WT and Cx43-Y286A is highly significant (P<0.0001).

four different experiments, and measured values were subjected to ANOVA statistical analysis. As shown in Fig. 3B, the degradation of Cx43-Y286A was significantly slower (P<0.0001) compared with that of Cx43-WT, with the half-life of Cx43 being increased from approximately 2 to 6 hours. The calculated rate of degradation for Cx43-WT (k=0.34 h^{-1}) was reduced 2.72-fold on mutation of Cx43-Y286A (k=0.125 h^{-1}). This approximates the increased steady-state abundance of Cx43-Y286A determined in our transient transfection assays (Fig. 2B; 3.5-fold), which can be expected if the accumulation rate is similar between wild-type and mutant protein. Collectively, these data show that the increase in the steady-state levels brought about by the substitution of tyrosine 286 with alanine results from a decrease in the rate of degradation of the mutant protein, suggesting that this amino acid plays an important role in Cx43 turnover.

Effects of proteasomal and lysosomal inhibitors on connexin turnover

Having established that the tyrosine 286 was involved in Cx43 stability, we next investigated the underlying cellular mechanism. As Cx43 has been shown to be degraded by both the proteasome and the lysosome (Beardslee et al., 1998; Laing and Beyer, 1995; Laing et al., 1997; Musil et al., 2000), we used pharmacological agents to inhibit the function of these intracellular compartments and measured Cx43 protein levels by western blot analysis (Fig. 4A). Note that because of the high connexin levels in Cx43-Y286A-expressing cells, reduced fluorography has been used on this blot for comparative analysis with the wildtype. Quantitative analysis of Cx43 expression with these treatments is shown in Fig. 4B. Treatment of Cx43-WT-expressing cells with the proteasomal inhibitor lactacystin (Fenteany et al., 1995) resulted in an approximate 3-fold increase only in the phosphorylated forms of Cx43 (Fig. 4B; right panel, grey columns). A similar effect could be observed using MG-132 (Palombella et al., 1994), an alternative proteasomal inhibitor (data not shown). By contrast, treatment of wild-type cells with two different lysosomal inhibitors, namely leupeptin and NH_4Cl (Hart et al., 1983), resulted in an increase in the nonphosphorylated form compared with the P1 and P2 forms, in the order of 1.5- and 2-fold, respectively (Fig. 4B; left panel, grey columns). These results are in agreement with the notion that both proteolytic pathways control the steady-state levels of Cx43.

In contrast to wild-type Cx43, Cx43-Y286A-expressing cells showed a very different response to proteasomal and lysosomal inhibitors (Fig. 4A). As previously shown, the ratio of P:NP forms of Cx43-WT (1:5.6) was already increased in untreated cells compared with Cx43-Y286A cells (1:2.6), and quantification further revealed that treatment with lactacystin had a much less prominent effect on the accumulation of the P1 and P2 forms (Fig. 4B; right panel, black columns). Most strikingly though, leupeptin and NH_4Cl no longer had an effect on the steady-state levels of unphosphorylated Cx43 (Fig. 4B; left panel, black columns). To further confirm this result, pulse-chase experiments were performed in the presence of NH_4Cl. As shown in Fig. 5A,B, treatment with NH_4Cl markedly reduced Cx43-WT turnover but had only a minor effect on the more stable mutant Cx43. Quantitative analysis of Cx43 expression levels revealed that the degradation rate of Cx43-WT (k=0.34 h^{-1}) was significantly reduced (~fivefold) in the presence of NH_4Cl (k=0.07h^{-1}), as indicated by ANOVA analysis (P<0.0001). But NH_4Cl treatment had a much weaker effect on the tyrosine mutant, as seen by the calculated degradation rates (control k=0.12 h^{-1}, NH_4Cl k=0.074 h^{-1}). In contrast to NH_4Cl, lactacystin did not affect the turnover of either Cx43-WT or Cx43-Y286A (data not shown; recall that only the NP form of Cx43 can be detected in the pulse-chase experiments). Taken together, these results show that the differences in Cx43 steady-state levels reflect a differential degradation of the wild-type and Y286A connexin primarily by the lysosome. Inhibition of the proteasome appears to affect mostly the ratio between NP and P forms and, to a much lesser extent, the total pool of Cx43.

Mutation of tyrosine 286 increases gap junctional staining

As a next step, we sought to examine the localisation of the wild-type and Y286A Cx43 within cells. Immunofluorescence localisation studies were therefore performed on both cell lines

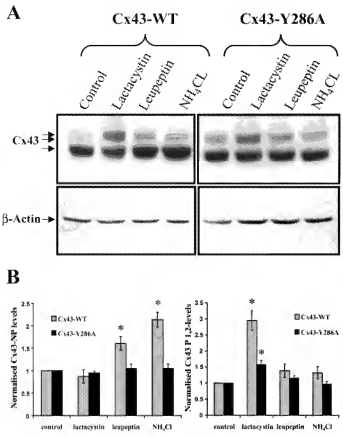


Fig. 4. Differential responses of wild-type Cx43 and Cx43-Y286A to both proteasomal and lysosomal inhibitors. (A) The stable cell lines, Cx43-WT and Cx43-Y286A, were treated for 3 hours with the proteasome inhibitor lactacystin (10 μM) or the lysosomal inhibitors leupeptin (10 μM) or NH_4Cl (10 mM). Western blot analysis was then performed on the cellular lysates using either an anti-Cx43 (top panel) or a β-actin (bottom panel) antibody. Steady-state levels of Cx43-Y286A were consistently greater than Cx43-WT, therefore reduced fluorography was performed for comparative analysis. (B) Quantitation of Cx43 protein levels in response to pharmacological agents. Experiments were performed as in A, and the amounts of Cx43-NP (left panel) and Cx43-P1,P2 (right panel) detected by western blot fluorography were quantified using a molecular imager FX. Mean ± s.e.m. values of normalised Cx43-WT (grey columns) and Cx43-Y286A (black columns) expression levels are shown (n=3 independent experiments). Asterisks represent P<0.01 versus control determined using the Student's t-test.

using an anti-Cx43 polyclonal antibody. Modest Cx43 immunoreactivity was detected in cells transfected with wild-type Cx43 (Fig. 6A; control), both intracellularly and at appositional membranes (indicated by arrows). The staining was strikingly stronger in the mutant cell line, which displayed larger and more abundant gap junctional plaques, as well as vesicular intracellular structures (Fig. 6B; control). We then examined the effects of proteasomal and lysosomal inhibitors on Cx43 localisation. Lactacystin, which specifically increased the phosphorylated forms of Cx43 (Fig. 2), markedly increased the gap junctional staining of Cx43-WT-expressing cells without any appreciable changes in intracellular labelling (Fig. 6A; lactacystin). This is in agreement with the notion that conversion of Cx43 to the P1 and P2 forms occurs after transport to the cell surface (Laird et al., 1995; Musil and Goodenough, 1991; Nagy et al., 1997). By contrast, lactacystin had no effect on the Cx43 labelling in the Y286A cell line (Fig. 6B).

Treatment of Cx43-WT-expressing cells with NH_4Cl gave a very different response than lactacystin. As can be seen in Fig. 6A, NH_4Cl, consistent with its inhibitory effect on lysosomal degradation, markedly increased Cx43 immunoreactivity within intracellular vesicles, which was probably due to the accumulation of undigested Cx43 in lysosomes. Moreover, the appositional staining was virtually abolished, suggesting that NH_4Cl may have some additional effects such as impairing the delivery or the recycling of connexins to the plasma membrane. In sharp contrast, NH_4Cl had no effect on Cx43 staining in the Cx43-Y286A mutant, with strong staining still clearly evident at cell-cell interfaces (Fig. 5B; NH_4Cl). The failure of NH_4Cl to disrupt the Cx43-Y286A-containing plaques may suggest that endocytosis had been affected in the connexin mutant. To further investigate this hypothesis, we used Brefeldin A (BFA), which prevents the delivery of newly synthesised proteins to the cell surface (Lippincott-Schwartz et al., 1991), to follow the fate of the Cx43 pool localised at the plasma membrane. In cells expressing wild-type Cx43, treatment with BFA resulted in a

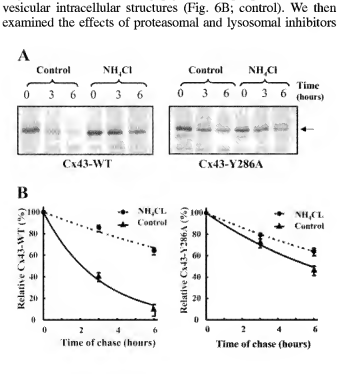


Fig. 5. The lysosomal inhibitor, NH_4Cl, affects wild-type, but less efficiently, mutant (Y286A) Cx43 turnover. (A) Pulse-chase analysis was performed on either Cx43-WT cells or Cx43-Y286A cells by pulsing for 60 minutes with [^{35}S]-methionine, and then chasing for 0, 3 or 6 hour periods in the absence (control) or presence of the lysosomal inhibitor, NH_4Cl. (B) Pulse-chase experiments were performed as in A, and the level of immunoprecipitated [^{35}S]-labelled Cx43 was quantified from fluorographs using a molecular imager FX. A plot of the mean±s.e.m. percentage of pulse-labelled Cx43 remaining after 0, 3 and 6 hours of chase (n=4 independent experiments), fitted using a mono-exponential decay function is shown. ANOVA statistical analysis reveals that the difference in the degradation curves between Cx43-WT (control and NH_4Cl) is highly significant (P<0.0001), in contrast to the Cx43-Y286A (control and NH_4Cl), which is significant (P=0.0303).

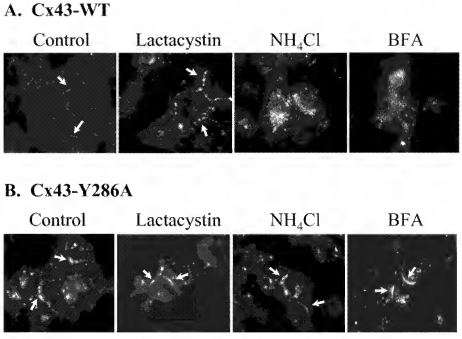
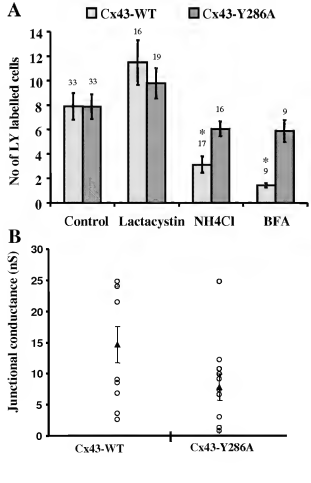


Fig. 6. Differential localisation of immunoreactive Cx43-WT compared with Cx43-Y286A in response to pharmacological agents. Cells expressing either wild-type Cx43 (A) or Y286A-mutated Cx43 (B) were incubated for 3 hours in new medium either alone (control) or supplemented with 10 μM lactacystin (Lactacystin), 10 mM NH_4Cl (NH_4Cl) or 2 μg/ml Brefeldin A (BFA). The cells were then fixed, immunostained with anti-Cx43 antibodies and visualised using fluorescence microscopy. Staining at appositional membranes is indicated by arrows.

loss of Cx43 immunoreactivity at appositional membranes while increasing the intracellular fluorescence. By contrast, in the Y286A mutant, large gap-junctional plaques were still clearly visible in the presence of BFA (Fig. 6B; BFA).



Tyrosine 286 affects Cx43 gap junctional communication

To investigate whether mutation within Cx43's tyrosine-based motif was also associated with a change in gap junctional communication (GJC), intracellular injections of lucifer yellow (LY) were carried out on both SKHep1 stable cell clones. As shown in Fig. 7A (control), both cell clones transferred LY to a similar extent, indicating that Cx43-Y286A can form functional gap junction channels. This was further confirmed by measuring junctional conductance in pairs of cells by the dual patch-clamp approach. No significant difference in electrical coupling between the wild-type and mutant cell clones was detected (Fig. 7B; P=0.064). We cannot exclude, however, the possibility that Cx43-Y286A has a lower permeability for LY and/or a reduced single-channel conductance that is compensated for by a larger number of channels at the cell surface.

Exposure of Cx43-WT and Cx43-Y286A cells to lactacystin

Fig. 7. NH_4Cl and BFA, but not lactacystin, affect wild-type, but not mutant (Y286A) Cx43-dependent gap junctional communication. (A) SKHep1 cells stably expressing Cx43-WT or Cx43-Y286A were incubated for 3 hours either in new medium alone (control) or new medium supplemented with 10 μM lactacystin (Lactacystin), 10 mM NH_4Cl (NH_4Cl) or 2 μg/ml Brefeldin A (BFA). Intercellular communication was then assessed by microinjecting individual cells within a cluster, and recording three minutes later the number of lucifer yellow (LY)-labelled cells by fluorescent microscopy, the results of which are expressed as mean±s.e.m. The number of injections is displayed above the column and asterisks indicate differences at P<0.01 versus control as determined by the Student's t-test. (B) Distribution of junctional conductance values (open circles) evaluated in Cx43-WT and Cx43-Y286A cell pairs monitored under the dual patch-clamp technique. The mean±s.e.m. junctional conductance value calculated for each distribution is indicated by the black triangle and error bars. No significant difference (P=0.064) between the two cell clones was detected using the Student's t-test for unpaired data.

did not significantly change the extent of dye coupling (Fig. 7A; lactacystin). By contrast, treatment with NH_4Cl had a marked effect on GJC. According to the immunohistochemistry, such a treatment should result in nearly complete loss of junctional staining in the wild-type, preventing LY transfer, while having little effect on the mutant. Indeed, under this condition, GJC in wild-type cells was significantly decreased compared with Y286A-cells, which remained LY competent (Fig. 7; NH_4Cl). Likewise, treatment with BFA had a marked effect on the ability of Cx43-WT cells to transfer LY, whereas diffusion of the fluorescent tracer remained efficient in cells expressing mutant Cx43 (Fig. 7; BFA).

Discussion

Data presented in this paper provide novel insights into understanding the process of endocytosis as an initial step in Cx43 gap junction channel degradation. Here we propose that a tyrosine-based sorting signal (YKLV) in the C-terminus of Cx43 controls gap junction turnover by affecting internalisation and targeting of Cx43 for degradation in the endosomal/lysosomal compartment.

Our data show that the tyrosine residue at position 286 (Y286) is a crucial amino acid involved in Cx43 turnover. Pulse-chase analysis in stably transfected SKHep1 cells indicates that substitution of Y286 with alanine increases the half-life of Cx43 from approximately 2 to 6 hours. Consistent with this, the steady-state levels of Cx43-Y286A is also elevated by a factor of ~3.5, which can be expected if one assumes equal biosynthesis rates of wild-type and mutant Cx43, which appears to be the case (Fig. 3A; compare time 0 of Cx43-WT with Cx43-Y286A). As outlined above, Y286 is part of two putative protein-protein interaction motifs (Fig. 1A): first, a PY-motif (Chen and Sudol, 1995), known to interact with WW domain-containing proteins such as the Nedd4/Nedd4-like family of ubiquitin-protein ligases (Rotin et al., 2000), and second, a tyrosine-based sorting signal (consensus: Yxxϕ, where ϕ is a hydrophobic amino acid), which is part of a family of degenerate motifs involved in the targeting of many transmembrane proteins to different cell compartments (Bonifacino and Dell'Angelica, 1999). Our mutational analysis of the region around Y286 and the subsequent transfection experiments in SKHep1 cells supports the notion that it is primarily the tyrosine motif that determines the stability of Cx43. Mutation of valine 289 to aspartate (V289D) in the tyrosine motif had the same drastic effect on the stability of Cx43 as Y286A, whereas mutation of proline 283 to leucine (P283L), which is part of the PY motif, increased Cx43 stability to a much weaker extent. Importantly, Y286A, being part of the two putative motifs, showed no additive effect on the stability (compared with V289D), as was the case for a double mutant (P283L/V289D) (data not shown). This therefore suggests that the PY motif plays only a limited, if any, role in the control of Cx43 turnover. It is possible, considering the steric properties of proline residues, that the P283L substitution has affected the functionality of the tyrosine-based sorting motif, providing an explanation for the comparably small increase in stability of this mutant. Indeed, it has been proposed recently that residues upstream of the crucial tyrosine may be important for the binding of effector molecules to tyrosine sorting motifs (Owen et al., 2001). Alternatively, we cannot exclude that the PY motif and the tyrosine motif act on the same pathway. The mutagenesis results further revealed that the substitution mutant Y286F behaves in a similar manner to that of the Y286A construct. This indicates that it is the tyrosine itself that is of crucial importance to Cx43 turnover and not its aromatic nature. Interestingly, tyrosine-based sorting signals are remarkably similar to the phospho-tyrosine-based motifs that direct SH2 domain binding (Pawson, 1995; Songyang et al., 1993). On this basis, we cannot exclude an involvement of tyrosine phosphorylation in this effect, especially as phosphorylation of such sorting signals has been shown to control protein localisation by regulating their interaction with the transport machinery (Bradshaw et al., 1997; Schaefer et al., 2002; Shiratori et al., 1997; Stephens and Banting, 1997). There is no evidence, however, that Y286 is a site of Cx43 phosphorylation.

Tyrosine-based sorting signals have been shown to interact with the medium chain (mu) subunits of the adaptor complexes (AP), which are components of the machinery involved in either clathrin-dependent or -independent formation of membrane-bound transport intermediates (e.g. coated vesicles) (Bonifacino and Dell'Angelica, 1999; Hirst et al., 1999; Hirst and Robinson, 1998; Simpson et al., 1996). Interestingly, annular gap junctions (e.g. internalised gap junctions) have been proposed to be clathrin coated (Larsen et al., 1979). Moreover, Cx43 has been shown to colocalise with clathrin (Huang et al., 1996), and has been found in close proximity to clathrin-coated pits within the plasma membrane (Naus et al., 1993), suggesting the possible involvement of a clathrin-mediated pathway in Cx43 trafficking. Therefore, it will be interesting to examine if the YKLV motif exerts its effects by interacting with one of the different adaptor complexes.

Several lines of evidence suggest that Y286 regulates the stability of Cx43 by controlling targeting of Cx43 for lysosomal degradation. Inhibitors of endosomal/lysosomal degradation, such as NH_4Cl and leupeptin, slowed down dramatically the degradation of Cx43-WT. This was accompanied by increased levels of the NP form of Cx43-WT (Fig. 4) and by an intracellular accumulation of the protein in vesicles that probably represent endosomes/lysosomes (Fig. 6). This is consistent with previous reports implicating the lysosomal system in the degradation of Cx43 (Berthoud, 2000; Laing and Beyer, 1995; Laing et al., 1997; Musil et al., 2000; Naus et al., 1993). By contrast, NH_4Cl (Fig. 4) and leupeptin (not shown) displayed only a marginal effect on the degradation rate and did not affect the steady-state levels of Cx43-Y286A, which remained elevated in all conditions (Fig. 4). The fact that NH_4Cl can still repress marginally the decay of the Y-mutant (Fig. 5) suggests that NH_4Cl has some small effects on the degradation mechanisms of Cx43 that are independent of Y286A. Brefeldin A, an inhibitor of transport from the ER to the plasma membrane, affected the cell-surface location and GJC of Cx43-WT, but not of Cx43-Y286A, indicating that mutation of tyrosine 286 may affect the retrieval of Cx43 from the plasma membrane. Our data further show that in transfected SKHep1 cells, the proteasome inhibitors lactacystin and MG-132 have only minor effects on overall stability (Fig. 4), but that they increase the Cx43 P forms in Cx43-WT-expressing cells, and, to a lesser extent, in Cx43-Y286A cells. Consistent with this observation, they also

increase the staining of Cx43-WT at appositional membranes (Fig. 5). Possibly, proteasome inhibitors may inhibit ER-dependent degradation (ERAD), as described previously for Cx32 and Cx43 (VanSlyke et al., 2000; VanSlyke and Musil, 2002), which could lead to an increased export of Cx43 to the cell surface. Alternatively, they may interfere with a direct role of the proteasome in the internalisation of Cx43, as has been described for the growth hormone receptor (van Kerkhof et al., 2000).

Surprisingly, the Y286A mutation increased the level of Cx43 immunoreactivity at cell-cell membrane contacts without affecting the extent of dye coupling. Dual patch-clamp analysis of pairs of Cx43-WT- and Cx43-Y286A-expressing cells confirmed that the two cell clones do not differ in terms of junctional conductance values. One explanation may be that the expression levels achieved in cells transfected with a CMV-driven Cx43 construct are high enough to cause maximal dye coupling that is already seen in Cx43-WT cells. Alternatively, the Cx43-Y286 mutant protein may be less efficient than Cx43-WT in transferring LY, which is compensated for by a larger number of channels at the cell surface. In our SKHep1 clones, no specific information could be obtained on the biophysical properties of the mutant Cx43 due to the high level of Cx45 channel activity in these cells (Moreno et al., 1995). Possibly, co-expression of Cx43 and Cx45 may form heteromeric channels with novel biophysical properties that may alter the normal behaviour of individual connexin components (Martinez et al., 2002). Despite these possibilities, mutation of Y286 prevented the decrease in GJC by inhibitors of the endosomal/lysosomal degradation pathways that was normally observed in cells expressing wild-type Cx43.

In conclusion, our data show that a tyrosine-based sorting signal present in the C-terminus of Cx43 controls turnover by targeting the protein for lysosomal degradation, thereby regulating the strength of gap junctional communication. The existence of such putative sequences in other connexin genes may suggest a common mechanism for the sorting of some members of the gap junction family. However, tyrosine-based signals are not the only recognised sequences to direct endocytosis and sorting of transmembrane proteins (Hu et al., 2001; Johnson and Kornfeld, 1992; Letourneur and Klausner, 1992; Stroh et al., 1999). Thus, the existence of several mechanisms for the sorting and degradation of gap junction channels made of distinct connexins may play important roles in various pathophysiological situations to maintain and/or modulate specific connexin expression and function at the junctional membrane.

We thank Jean-Daniel Horisberger, Hugues Abriel, Dario Diviani, Laura Stanasila and Christophe Debonneville for critically reading the manuscript, Hugues Abriel and Laura Stanasila for help with statistical analysis, as well as Melete Solomon and Tecla Dudez for technical help. This work was supported by grants from the Leenaards Foundation (to M.C., N.D. and O.S.), Vaincre la Mucoviscidose and the Swiss National Science Foundation #3100-067120.01 (to M.C.) and #31-64052.00 (to O.S.).

References

Abriel, H., Loffing, J., Rebhun, J. F., Pratt, J. H., Horisberger, J.-D., Rotin, D. and Staub, O. (1999). Defective regulation of the epithelial Na^+ channel (ENaC) by Nedd4 in Liddle's syndrome. *J. Clin. Invest.* **103**, 667-673.

Abriel, H., Kamynina, E., Horisberger, J.-D. and Staub, O. (2000). Regulation of the cardiac voltage-gated Na^+ channel (rH1) by the ubiquitin-protein ligase Nedd4. *FEBS Lett.* **466**, 377-380.

Beardsley, M. A., Laing, G., Beyer, E. C. and Saffitz, J. E. (1998). Rapid turnover of connexin43 in the adult rat heart. *Circ. Res.* **83**, 629-635.

Berthoud, V. M., Tadros, P. N. and Beyer, E. C. (2000). Connexin and gap junction degradation. *Methods* **20**, 180-187.

Bonifacino, J. S. and Dell'Angelica, E. C. (1999). Molecular bases for the recognition of tyrosine-based sorting signals. *J. Cell Biol.* **145**, 923-926.

Bradshaw, J. D., Lu, P., Leytze, G., Rodgers, J., Schieven, G. L., Bennett, K. L., Linsley, P. S. and Kurtz, S. E. (1997). Interaction of the cytoplasmic tail of CTLA-4 (CD152) with a clathrin-associated protein is negatively regulated by tyrosine phosphorylation. *Biochemistry* **36**, 15975-15982.

Chen, H. I. and Sudol, M. (1995). The WW domain of Yes-associated protein binds a novel proline-rich ligand that differs from the consensus established for SH3-binding modules. *Proc. Natl. Acad. Sci. USA* **92**, 7819-7823.

Darrow, B. J., Laing, J. G., Lampe, P. D., Saffitz, J. E. and Beyer, E. C. (1995). Expression of multiple connexins in cultured neonatal rat ventricular myocytes. *Circ. Res.* **76**, 381-387.

Fenteany, G., Standaert, R. F., Lane, W. S., Chois, S., Corey, E. J. and Schreiber, S. L. (1995). Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification of lactacystin. *Science* **268**, 726-731.

Gaietta, G., Deerinck, T. J., Adams, S. R., Bouwer, J., Tour, O., Laird, D. W., Sosinsky, G. E., Tsien, R. Y. and Ellisman, M. H. (2002). Multicolor and electron microscopic imaging of connexin trafficking. *Science* **296**, 503-507.

Goodenough, D. A., Goliger, J. A. and Paul, D. L. (1996). Connexins, connexons, and intercellular communication. *Annu. Rev. Biochem.* **65**, 475-502.

Hart, P. D., Young, M. R., Jordan, M. M., Perkins, W. J. and Geisow, M. J. (1983). Chemical inhibitors of phagosome-lysosome fusion in cultured macrophages also inhibit saltatory lysosomal movements. A combined microscopic and computer study. *J. Exp. Med.* **158**, 477-492.

Hirst, J., Bright, N. A., Rous, B. and Robinson, M. S. (1999). Characterization of a fourth adaptor-related protein complex. *Mol. Biol. Cell* **10**, 2787-2802.

Hirst, J. and Robinson, M. S. (1998). Clathrin and adaptors. *Biochim. Biophys. Acta* **1404**, 173-193.

Hu, W., Howard, M. and Lukacs, G. L. (2001). Multiple endocytic signals in the C-terminal tail of the cystic fibrosis transmembrane conductance regulator. *Biochem. J.* **354**, 561-572.

Huang, X. D., Horackova, M. and Pressler, M. L. (1996). Changes in the expression and distribution of connexin 43 in isolated cultured adult guinea pig cardiomyocytes. *Exp. Cell Res.* **228**, 254-261.

Johnson, K. F. and Kornfeld, S. (1992). A His-Leu-Leu sequence near the carboxyl terminus of the cytoplasmic domain of the cation-dependent mannose 6-phosphate receptor is necessary for the lysosomal enzyme sorting function. *J. Biol. Chem.* **267**, 17110-17115.

Jordan, K., Chodock, R., Hand, A. R. and Laird, D. W. (2001). The origin of annular junctions: a mechanism of gap junction internalization. *J. Cell Sci.* **114**, 763-773.

Kirchhausen, T., Bonifacino, J. S. and Riezman, H. (1997). Linking cargo to vesicle formation: receptor tail interactions with coat proteins. *Curr. Opin. Cell Biol.* **9**, 488-495.

Kumar, N. M. and Gilula, N. B. (1996). The gap junction communication channel. *Cell* **84**, 381-388.

Kwak, B. R., Saez, J. C., Wilders, R., Chanson, M., Fishman, G. I., Hertzberg, E. L., Spray, D. C. and Jongsma, H. J. (1995). Effects of cGMP-dependent phosphorylation on rat and human connexin43 gap junction channels. *Pflugers Arch.* **430**, 770-778.

Laing, J. G. and Beyer, E. C. (1995). The gap junction protein connexin43 is degraded via the ubiquitin proteasome pathway. *J. Biol. Chem.* **270**, 26399-26403.

Laing, J. G., Tadros, P. N., Westphale, E. M. and Beyer, E. C. (1997). Degradation of connexin43 gap junctions involves both the proteasome and the lysosome. *Exp. Cell Res.* **236**, 482-492.

Laird, D. W. (1996). The life cycle of a connexin: gap junction formation, removal, and degradation. *J. Bioenerg. Biomembr.* **28**, 311-318.

Laird, D. W., Puranam, K. L. and Revel, J. P. (1991). Turnover and phosphorylation dynamics of connexin43 gap junction protein in cultured cardiac myocytes. *Biochem. J.* **273**, 67-72.

Laird, D. W., Castillo, M. and Kasprzak, L. (1995). Gap junction turnover, intracellular trafficking, and phosphorylation of connexin43 in brefeldin A-treated rat mammary tumor cells. *J. Cell Biol.* **131**, 1193-1203.

Lampe, P. D. (1994). Analyzing phorbol ester effects on gap junctional communication: a dramatic inhibition of assembly. *J. Cell Biol.* **127**, 1895-1905.

Larsen, W. J. and Hai, N. (1978). Origin and fate of cytoplasmic gap junctional vesicles in rabbit granulosa cells. *Tissue Cell* **10**, 585-598.

Larsen, W. J., Tung, H. N., Murray, S. A. and Swenson, C. A. (1979). Evidence for the participation of actin microfilaments and bristle coats in the internalization of gap junction membrane. *J. Cell Biol.* **83**, 576-587.

Letourneur, F. and Klausner, R. D. (1992). A novel di-leucine motif and a tyrosine-based motif independently mediate lysosomal targeting and endocytosis of CD3 chains. *Cell* **69**, 1143-1157.

Lippincott-Schwartz, J., Yuan, L., Tipper, C., Amherdt, M., Orci, L. and Klausner, R. D. (1991). Brefeldin A's effects on endosomes, lysosomes, and the TGN suggest a general mechanism for regulating organelle structure and membrane traffic. *Cell* **67**, 601-616.

Luke, R. A. and Saffitz, J. E. (1991). Remodeling of ventricular conduction pathways in healed canine infarct border zones. *J. Clin. Invest.* **87**, 1594-1602.

Martinez, A. D., Hayrapetyan, V., Moreno, A. P. and Beyer, E. C. (2002). Connexin43 and connexin45 form heteromeric gap junction channels in which individual components determine permeability and regulation. *Circ. Res.* **90**, 1100-1107.

Moreno, A. P., Saez, J. C., Fishman, G. I. and Spray, D. C. (1994). Human connexin43 gap junction channels. Regulation of unitary conductances by phosphorylation. *Circ. Res.* **74**, 1050-1057.

Moreno, A. P., Laing, J. G., Beyer, E. C. and Spray, D. C. (1995). Properties of gap junction channels formed of connexin45 endogenously expressed in human hepatoma (SKHep1) cells. *Am. J. Physiol.* **268**, C356-C365.

Musil, L. S. and Goodenough, D. A. (1991). Biochemical analysis of connexin43 intracellular transport, phosphorylation, and assembly into gap junctional plaques. *J. Cell Biol.* **115**, 1357-1374.

Musil, L. S., Beyer, E. C. and Goodenough, D. A. (1990a). Expression of the gap junction protein connexin43 in embryonic chick lens: molecular cloning, ultrastructural localization, and post- translational phosphorylation. *J. Membr. Biol.* **116**, 163-175.

Musil, L. S., Cunningham, B. A., Edelman, G. M. and Goodenough, D. A. (1990b). Differential phosphorylation of the gap junction protein connexin43 in junctional communication-competent and -deficient cell lines. *J. Cell Biol.* **111**, 2077-2088.

Musil, L. S., Le, A. C., VanSlyke, J. K. and Roberts, L. M. (2000). Regulation of connexin degradation as a mechanism to increase gap junction assembly and function. *J. Biol. Chem.* **275**, 25207-25215.

Nagy, J. I., Li, W. E., Roy, C., Doble, B. W., Gilchrist, J. S., Kardami, E. and Hertzberg, E. L. (1997). Selective monoclonal antibody recognition and cellular localization of an unphosphorylated form of connexin43. *Exp. Cell Res.* **236**, 127-136.

Naus, C. C., Hearn, S., Zhu, D., Nicholson, B. J. and Shivers, R. R. (1993). Ultrastructural analysis of gap junctions in C6 glioma cells transfected with connexin43 cDNA. *Exp. Cell Res.* **206**, 72-84.

Nelson, R. M. and Long, G. L. (1989). A general method of site-specific mutagenesis using a modification of the *Thermus aquaticus* polymerase chain reaction. *Anal. Biochem.* **180**, 147-151.

Ohno, H., Stewart, J., Fournier, M. C., Bosshart, H., Rhee, I., Miyatake, S., Saito, T., Gallusser, A., Kirchhausen, T. and Bonifacino, J. S. (1995). Interaction of tyrosine-based sorting signals with clathrin-associated proteins. *Science* **269**, 1872-1875.

Owen, D. J. and Evans, P. R. (1998). A structural explanation for the recognition of tyrosine-based endocytotic signals. *Science* **282**, 1327-1332.

Owen, D. J., Setiadi, H., Evans, P. R., McEver, R. P. and Green, S. A. (2001). A third specificity-determining site in mu 2 adaptin for sequences upstream of Yxx phi sorting motifs. *Traffic* **2**, 105-110.

Palombella, V., Rando, O., Goldberg, A. L. and Maniatis, T. (1994). Ubiquitin and the proteasome are required for processing the NF-κB1 precursor and the activation of NF-κB. *Cell* **78**, 773-785.

Pawson, T. (1995). Protein modules and signalling networks. *Nature* **373**, 573-580.

Rabionet, R., Lopez-Bigas, N., Arbones, M. L. and Estivill, X. (2002). Connexin mutations in hearing loss, dermatological and neurological disorders. *Trends Mol. Med.* **8**, 205-212.

Rogers, S., Wells, R. and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-368.

Rotin, D., Staub, O. and Haguenauer-Tsapis, R. (2000). Ubiquitination and endocytosis of plasma membrane proteins: role of Nedd4/Rsp5p family of ubiquitin-protein ligases. *J. Membr. Biol.* **176**, 1-17.

Rutz, M. L. and Huber, D. F. (2001). Supramolecular dynamics of gap junctions. *Eur. J. Cell Biol.* **80**, 20-30.

Schaefer, A. W., Kamei, Y., Kamiguchi, H., Wong, E. V., Rapoport, I., Kirchhausen, T., Beach, C. M., Landreth, G., Lemmon, S. K. and Lemmon, V. (2002). L1 endocytosis is controlled by a phosphorylation-dephosphorylation cycle stimulated by outside-in signaling by L1. *J. Cell Biol.* **157**, 1223-1232.

Schwake, M., Friedrich, T. and Jentsch, T. J. (2001). An internalization signal in ClC-5, an endosomal Cl-channel mutated in dent's disease. *J. Biol. Chem.* **276**, 12049-12054.

Shiratori, T., Miyatake, S., Ohno, H., Nakaseko, C., Isono, K., Bonifacino, J. S. and Saito, T. (1997). Tyrosine phosphorylation controls internalization of CTLA-4 by regulating its interaction with clathrin-associated adaptor complex AP- 2. *Immunity* **6**, 583-589.

Simon, A. M. and Goodenough, D. A. (1998). Diverse functions of vertebrate gap junctions. *Trends Cell Biol.* **8**, 477-483.

Simpson, F., Bright, N. A., West, M. A., Newman, L. S., Darnell, R. B. and Robinson, M. S. (1996). A novel adaptor-related protein complex. *J. Cell Biol.* **133**, 749-760.

Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J. et al. (1993). SH2 domains recognize specific phosphopeptide sequences. *Cell* **72**, 767-778.

Spray, D. C. (1996). Molecular physiology of gap junction channels. *Clin. Exp. Pharmacol. Physiol.* **23**, 1038-1040.

Staub, O., Dho, S., Henry, P. C., Correa, J., Ishikawa, T., McGlade, J. and Rotin, D. (1996). WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na+ channel deleted in Liddle's syndrome. *EMBO J.* **15**, 2371-2380.

Stephens, D. J. and Banting, G. (1997). Insulin dependent tyrosine phosphorylation of the tyrosine internalisation motif of TGN38 creates a specific SH2 domain binding site. *FEBS Lett.* **416**, 27-29.

Stroh, A., Schafer, W., Berghofer, S., Eickmann, M., Teuchert, M., Burger, I., Klenk, H. D. and Garten, W. (1999). A mono phenylalanine-based motif (F790) and a leucine-dependent motif (LI760) mediate internalization of furin. *Eur. J. Cell Biol.* **78**, 151-160.

Traub, O., Druge, P. M. and Willecke, K. (1983). Degradation and resynthesis of gap junction protein in plasma membranes of regenerating liver after partial hepatectomy or cholestasis. *Proc. Natl. Acad. Sci. USA* **80**, 755-759.

van Kerkhof, P., Govers, R., Alves dos Santos, C. M. and Strous, G. J. (2000). Endocytosis and degradation of the growth hormone receptor are proteasome-dependent. *J. Biol. Chem.* **275**, 1575-1580.

VanSlyke, J. K. and Musil, L. S. (2002). Dislocation and degradation from the ER are regulated by cytosolic stress. *J. Cell Biol.* **157**, 381-394.

VanSlyke, J. K., Deschenes, S. M. and Musil, L. S. (2000). Intracellular transport, assembly, and degradation of wild-type and disease-linked mutant gap junction proteins. *Mol. Biol. Cell* **11**, 1933-1946.

Varshavsky, A. (1992). The N-end rule. *Cell* **69**, 725-735.

Vaughan, D. K. and Lasater, E. M. (1990). Renewal of electrotonic synapses in teleost retinal horizontal cells. *J. Comp. Neurol.* **299**, 364-374.

Yamasaki, H. and Naus, C. C. (1996). Role of connexin genes in growth control. *Carcinogenesis* **17**, 1199-1213.

Communication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 265, No. 10, Issue of April 5, pp. 5349-5352, 1990

Printed in U.S.A.

A New Class of Lysosomal/ Vacuolar Protein Sorting Signals*

(Received for publication, December 11, 1989)

Daniel J. Klionsky‡ and Scott D. Emr§

From the Division of Biology, California Institute of Technology, Pasadena, California 91125

A number of inherited lysosomal diseases are known to result from missorting of lysosomal proteins. Considerable attention has been directed toward an understanding of this sorting pathway, and it has become apparent that different mechanisms are used for the sorting of lysosomal membrane and soluble proteins. Protein sorting to the yeast vacuole/lysosome provides a simple model system to study this process. We have mapped the first sorting signal in a vacuolar membrane protein, repressible alkaline phosphatase, and have shown it to be both necessary and sufficient for vacuolar delivery of this enzyme. The sorting information is confined to the transmembrane and cytoplasmic tail region of this type II integral membrane protein. The location of this sorting signal provides an explanation for some of the differences observed between membrane and soluble vacuolar protein sorting.

Normal eukaryotic cell function depends on the separation of competing chemical reactions into distinct subcellular compartments. Constituent proteins of each organelle must be efficiently delivered from their site of synthesis in the cytoplasm to the appropriate subcellular compartment where they execute their function. Defects in intracellular protein sorting are known to result in a variety of human disorders (1–3). The secretory pathway (4) has long been used as a model system to study the processes involved in protein sorting, and much of this work has focused on the delivery of proteins to the mammalian lysosome (5). Recently, work from many laboratories has shown that similar sorting processes take place in the yeast *Saccharomyces cerevisiae* (6). An analysis of protein sorting to the yeast vacuole, which is analogous to the mammalian lysosome, has been useful in identifying both the *cis*-acting signals involved in targeting proteins to the vacuole and the *trans*-acting components of the machinery that mediate the sorting process (7–11).

Lysosomal membrane proteins are targeted by a mechanism which is in part different from that used by soluble lysosomal hydrolases. The mannose 6-phosphate recognition system that participates in the sorting of many soluble lysosomal enzymes, for example, is not used for at least some lysosomal membrane proteins (12–14). The underlying factors which allow these two classes of proteins to be differentially sorted, however, are not understood. We have recently shown that yeast-repressible alkaline phosphatase (ALP),[1] the product of the *PHO8* gene, is a type II integral vacuolar membrane protein. Its delivery to the vacuole exhibits characteristics different from those of soluble vacuolar hydrolases (15). Here we report on the identification of the vacuolar sorting signal contained within ALP. Its location in ALP helps to provide an explanation for the observed differences between soluble and membrane vacuolar protein sorting.

EXPERIMENTAL PROCEDURES

Materials—Lyticase was obtained from Enzogenetics (Corvallis, OR), Trans^{35}S-label was from ICN Radiochemicals (Irvine, CA), α_2-macroglobulin was from Boehringer Mannheim, and Autofluor was from National Diagnostics. All other chemicals and reagents were of a commercially available high purity. Antisera to alkaline phosphatase and invertase were prepared as described previously (9, 15). Antiserum to α-1,3-mannose linkages was generously provided by R. Schekman (16).

Strains—Yeast strains used were SEY2108 *MATa*, *ura3-52*, *leu2-3,112*, *suc2-Δ9*, Δ*prc1*::*LEU2* (17); SEY6211 *MATa*, *ura3-52*, *leu2-3,112*, *his3-Δ200*, *trp1-Δ901*, *ade2-101*, *suc2-Δ9*, *GAL* (10); and DKY6280 *MATa*, *ura3-52*, *leu2,3-112*, *his3-Δ200*, *trp1-Δ901*, *ade2-101*, *suc2-Δ9*, Δ*pho8*::*TRP1* (15).

Plasmid Constructions—The ALP-Inv19 fusion was constructed using the naturally occurring *Xho*I site in *PHO8* by cloning an *Eco*RI-*Xho*I fragment into the *Eco*RI-*Sal*I sites of the *SUC2* fusion vector pSEYC308 (9). The ALP-Inv191 fusion utilized a *Bgl*II site in *PHO8* and was cloned as an *Eco*RI-*Bgl*II fragment into the *Eco*RI-*Bam*HI sites of plasmid pSEYC306 (7). The ALP-Inv52, -67, and -77 gene fusions were constructed by inserting a *Bam*HI restriction site after the indicated amino acid codon by oligonucleotide-directed mutagenesis (7). Each was then cloned as an *Eco*RI-*Bam*HI fragment into pSEYC306. The plasmid pDKY1 (CPY20/ALPΔ1-53) was constructed as follows: *PHO8* was mutagenized as above to insert a *Bam*HI site after amino acid 52. The resulting ~2.5-kilobase *Bam*HI fragment encoding amino acids 54–566 of ALP was isolated and cloned into the *Bam*HI site of pCY1-20 which contains the first 20 amino acid residues of preprocarboxypeptidase Y (7). The presence of the *Bam*HI site resulted in the introduction of an additional amino acid residue (aspartic acid) at position 21 of the hybrid protein.

Immunoprecipitation—Yeast strain SEY2108 expressing the individual hybrid proteins was grown in YNB glucose (yeast nitrogen base containing 2% glucose). Two OD_{600} units of cells were collected by centrifugation and resuspended in YNB glucose at a concentration of 4 absorbance units/ml. Tunicamycin (final concentration of 20 μg/ml) was added to inhibit *N*-linked glycosylation 15 min before the addition of Trans^{35}S-label. Labeling was allowed to continue for 45 min followed by a 15-min chase in the presence of cold methionine and cysteine (10 mM). Invertase antiserum was used to precipitate the hybrid proteins as described previously (9). Radiolabeled proteins were electrophoresed on an SDS-polyacrylamide gel. After electrophoresis, gels were fixed in acetic acid (10%) and treated with Autofluor.

To examine the CPY20/ALPΔ1-53 hybrid protein, yeast strain DKY6280 harboring plasmid pDKY1 was converted to spheroplasts and labeled as described previously (10). Spheroplasts were labeled for 7.5 min, chased, and separated into intracellular (spheroplast) and extracellular (periplasm and medium) fractions and immunoprecipitated twice with antiserum to ALP. To determine the presence of Golgi-specific carbohydrate modifications, the 30-min chase sample was divided in half after the first precipitation with ALP antiserum. The second immunoprecipitation was carried out on one half of the

* This work was supported by a National Science Foundation Presidential Young Investigator award (to S. D. E.) and research fellowships from the Helen Hay Whitney Foundation and the American Cancer Society, California Division (to D. J. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Dept. of Microbiology, University of California, Davis, CA 95616.

§ To whom correspondence should be sent.

[1] The abbreviations used are: ALP, alkaline phosphatase; SDS, sodium dodecyl sulfate; ER, endoplasmic reticulum.

sample using ALP antiserum and on the remaining half using α-1,3-mannose antiserum.

Vacuole Isolation—Yeast cells expressing the various hybrid proteins were converted to spheroplasts, gently lysed with DEAE-dextran, and fractionated on Ficoll discontinuous density gradients as described previously (9, 17). The isolated fractions were assayed along with the crude cell extract for the following enzyme activities: α-mannosidase (vacuole membrane marker), invertase (hybrid protein), NADPH cytochrome *c* reductase (ER membrane marker), and α-glucosidase (cytoplasmic marker). Vacuole yields were typically 25–30% and exhibited less than 3% contamination with marker enzymes from other cellular compartments.

RESULTS AND DISCUSSION

A gene fusion approach has proven very useful for the identification of vacuolar sorting signals in soluble vacuolar hydrolases (7–9). We have extended this approach to study the vacuolar membrane protein ALP. The yeast *SUC2* gene codes for the enzyme invertase, which is normally localized to the periplasm. Invertase participates in sucrose metabolism by hydrolyzing sucrose to glucose and fructose, and simple biochemical assays exist for detecting its activity. Plasmid vectors have been described which contain a slightly truncated *SUC2* gene, lacking the coding sequence for the NH_2-terminal signal peptide and the first two amino acids of the mature protein (7, 9). We used these vectors to construct five *PHO8-SUC2* gene fusions containing different sized amino-terminal coding segments from the *PHO8* gene fused to a constant fragment of the *SUC2* gene (Fig. 1). Each of these gene fusions directs the synthesis of ALP-invertase hybrid proteins that exhibit invertase activity in yeast. The apparent molecular weight of each of the hybrid proteins synthesized in the presence of tunicamycin, which blocks the addition of *N*-linked oligosaccharides, is consistent with that expected from the deduced amino acid sequence of each of the *PHO8-SUC2* gene fusion constructs (Fig. 2). The ALP-Inv hybrid proteins contain all the oligosaccharide addition sites present in wild-type invertase, but they lack the two sites present in wild-type ALP. Invertase normally undergoes extensive glycosyl processing in the Golgi complex resulting in a hyperglycosylated protein with extensive outer chain carbohydrates (18, 19). The hyperglycosylated invertase migrates as a high molecular weight smear on SDS-polyacrylamide gels due to the variable length of the outer mannose chains that are added to the core oligosaccharides. When the ALP-Inv hybrid proteins are immunoprecipitated in the absence of tunicamycin, all but one show the migration pattern typically seen with wild-type invertase (data not shown). The ALP-Inv19 hybrid protein migrates as a discrete band, and its migration position is the same in the presence or absence of tunicamycin. ALP-Inv19 is apparently not glycosylated suggesting that it is not translocated into the ER lumen, the site of core oligosaccharide addition (20). This was expected because an analysis of the deduced amino acid sequence of ALP (21) indicates that the extreme NH_2 terminus of the protein does not contain a sequence which fits with the normal hydrophobic consensus for signal peptides (15, 22). Since the hybrid proteins containing 52 or more amino acids from the NH_2 terminus of ALP are substrates for *N*-linked glycosylation, they must be translocated across the ER membrane. The only sequence sufficiently hydrophobic to function as a signal peptide in these hybrid proteins, or in wild-type ALP, is located near the NH_2 terminus and spans amino acid residues 34–59. This part of the protein presumably functions as an internal uncleaved signal sequence, in addition to its role in anchoring the protein in the membrane (15). Hybrid proteins containing the NH_2-terminal hydrophobic domain, including the ALP-Inv52 hybrid protein, maintain the same orientation in the membrane as wild-type ALP, based on sensitivity to exogenous proteases (15) (data not shown). The presence of invertase in the fusion constructions therefore does not interfere with normal membrane interaction of the ALP hydrophobic domain.

The invertase activity associated with each of the hybrid proteins serves as a biochemical marker to identify each hybrid protein's location within the cell. To accurately determine the location of the invertase activity, we fractionated cells containing the various hybrid proteins on Ficoll step density gradients (7). Purified vacuoles were isolated from these gradients and assayed along with the crude cell extract to determine the percentage of the hybrid protein (invertase activity) that had been localized to the vacuole (Fig. 1). The invertase activity associated with the ALP-Inv19 fusion does not cofractionate with the purified vacuoles but remains in the soluble cell fraction (Fig. 1). This is consistent with the glycosylation results and confirmed that the ALP-Inv19 hybrid protein does not contain a signal peptide allowing for ER translocation. ALP-Inv fusions containing 52, 67, 77, or 191 amino acids from ALP showed nearly quantitative cofractionation of the invertase activity with α-mannosidase activity, a marker enzyme of the vacuole. Consistent with these fraction-

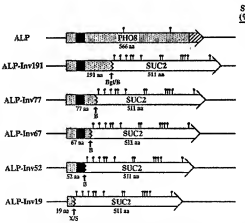


FIG. 1. ***PHO8-SUC2* gene fusions.** The *PHO8* gene is divided into the coding regions for the propeptide (*hatched box*), the hydrophobic domain (*solid box*), and the additional mature enzyme (*stippled box*). The *SUC2* gene is indicated by the open box. The approximate positions of the two asparagine-linked core oligosaccharides as determined from the deduced amino acid sequence of the ALP protein (19) and the 13 potential sites for core oligosaccharide addition on invertase are indicated above the *PHO8* and *SUC2* genes, respectively. Cofractionation of the *PHO8* gene product and the hybrid protein products of each of the *PHO8-SUC2* gene fusions with isolated vacuoles is as indicated; the invertase activity is expressed as the percent cofractionation with the vacuole marker α-mannosidase in isolated vacuoles. *aa*, amino acids; *B*, *Bam*HI; *X*, *Xho*I; *S*, *Sal*I; *Bgl*, *Bgl*II.

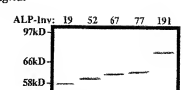


FIG. 2. **Immunoprecipitation of unglycosylated ALP-Inv hybrid proteins.** Labeling of hybrid proteins in strain SEY2108 in the presence of tunicamycin and polyacrylamide gel electrophoresis was carried out as described under "Experimental Procedures." The number of amino acids of pro-ALP fused to invertase is indicated above each *lane*. The positions of protein standards (molecular weight, 10^3) are indicated.

ation data, yeast cells deleted for the chromosomal *SUC* loci and harboring these *PHO8-SUC2* fusion plasmids are phenotypically Suc^-. These results indicate that as few as 52 NH_2-terminal amino acids of ALP contain sufficient information to direct invertase to the yeast vacuole. Since the ALP-Inv hybrid proteins are delivered to the vacuole even though they do not contain the oligosaccharide addition sites present in wild-type ALP, it is unlikely that these carbohydrates contribute to the vacuolar sorting of the wild-type protein. This is in agreement with earlier observations showing that alkaline phosphatase activity is localized to the vacuole in the presence of tunicamycin (23).

The fusion results indicate that the NH_2 terminus of ALP, including the cytoplasmic tail and/or hydrophobic domain, contains sorting information which is sufficient to direct a normally secreted enzyme to the vacuole. To confirm the role of this amino-terminal portion of ALP in directing the vacuolar localization of the wild-type protein, we deleted this region from an otherwise wild-type ALP protein and examined the effect of the deletion on the sorting of ALP. Because this NH_2-terminal sequence also functions as a signal peptide for translocation into the ER lumen, however, it was necessary to supply an exogenous signal peptide. For this reason, we fused the first 20 amino acid residues of preprocarboxypeptidase Y, which constitute the transient signal peptide of carboxypeptidase Y, in front of an ALP protein that had been deleted for its first 53 residues (CPY20/ALPΔ1-53, Fig. 3). Earlier work had shown that this part of preprocarboxypeptidase Y functions as an efficient signal to direct ER translocation but is processed from the protein in the ER and does not contain any vacuolar targeting information (7). A plasmid encoding this hybrid construct, pDKY1, was transformed into strain DKY6280 (15), which is deleted for the chromosomal *PHO8* locus. A pulse-chase analysis was used to examine the effect of the deletion on the sorting of ALP (Fig. 3). We found that deletion of the cytoplasmic tail and most of the hydrophobic domain results in mislocalization of the mutated ALP protein. All of the mutant ALP protein that is transferred from the ER to the Golgi complex as monitored by acquisition of a Golgi-specific glycosyl modification is subsequently missorted and secreted from the cell. This glycosyl modification, the addition of particular α-1,3-mannose linkages, is indicative of a protein having had access to certain of the Golgi mannosyltransferases (16). All of the protein which is immunoprecipitated by an α-1,3-mannose-specific antiserum is secreted from the cell by the 90-min chase point. The remainder of the protein (approximately 50%) could not be precipitated by the α-1,3-mannose antiserum and presumably accumulates in the ER, possibly as the result of improper folding (24). It is not likely that ER accumulation is due to retention of the carboxypeptidase Y signal peptide because: 1) the fusion joint between CPY20 and ALPΔ1-53 was constructed so as to maximize the probability of signal peptide cleavage from the hybrid protein based on the rules established by von Heijne ((22), S value 4.4 for preprocarboxypeptidase Y and 5.6 for the hybrid protein); 2) the protein migrates on an SDS-polyacrylamide gel with a molecular weight consistent with removal of the signal peptide; and 3) the protein that accumulates in the ER can be released from the ER lumen by treatment with 0.2% saponin (data not shown), suggesting it is not anchored in the ER membrane. In contrast, wild-type proALP maintains its association with the ER membrane in the presence of saponin (15). The ER and Golgi forms of CPY20/ALPΔ1-53 are resolved more easily by SDS-polyacrylamide gel electrophoresis than seen previously for wild-type ALP (15). This may be explained by the reduced molecular weight of this construct or the deletion of the hydrophobic domain which could cause aberrant migration. These results show that the NH_2 terminus of ALP is not only sufficient to divert a normally secreted enzyme to the vacuole but is also necessary for correct delivery of the wild-type protein to the vacuole.

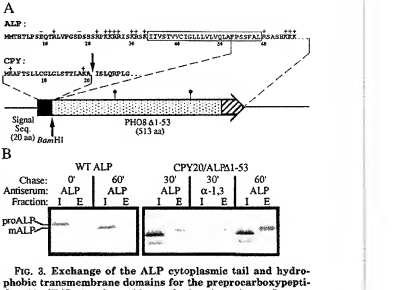


FIG. 3. **Exchange of the ALP cytoplasmic tail and hydrophobic transmembrane domains for the preprocarboxypeptidase Y (*CPY*) signal peptide results in missorting and secretion of ALP.** *A*, the sequence of the NH_2-terminal 67 amino acids (*aa*) of ALP and 28 amino acids of preprocarboxypeptidase Y. *Dotted lines* indicate the construction of the CPY20/ALPΔ1-53 hybrid protein, which is described under "Experimental Procedures." The *arrow* above the sequence marks the site of signal peptide processing (29). *B*, immunoprecipitation of wild-type (*WT*) ALP (strain SEY6211) and CPY20/ALPΔ1-53. Spheroplasts of strain SEY6211 or of strain DKY6280 harboring plasmid pDKY1 were labeled for 7.5 min, chased for the indicated times, and separated into intracellular and extracellular fractions. The samples were immunoprecipitated with ALP antiserum followed by a second immunoprecipitation with antiserum to either ALP or α-1,3-mannose linkages, as indicated. The positions of wild-type mature and precursor ALP are marked. The lowest molecular weight band in the CPY20/ALPΔ1-53 immunoprecipitation (*lanes 1* and *5* of the *second gel panel*) migrates with an apparent molecular weight predicted for the unglycosylated hybrid protein and presumably represents a small amount of hybrid protein that has not been translocated into the ER. *I*, intracellular (spheroplast); *E*, extracellular (periplasm and medium).

Our finding that the vacuolar sorting signal in ALP is located in the cytoplasmic tail and/or hydrophobic domain is interesting in regard to our earlier observation that ALP is not missorted in the presence of bafilomycin (15), a specific and potent inhibitor of the vacuolar ATPase (25). Bafilomycin causes an increase in the pH of the vacuole, and possibly other compartments, and results in extensive missorting of the soluble vacuolar hydrolases proteinase A/proteinase B, and carboxypeptidase Y but has no effect on the processing of ALP (15, 26). ALP presumably interacts with sorting components that are in the cytoplasm or are within the lipid bilayer while the soluble hydrolases must interact with lumenal sorting components. Obviously these latter interactions are expected to be more susceptible to changes in the chemical environment of the vacuole lumen. This difference in the spatial location of the vacuolar sorting signal provides a simple explanation for the differences seen in sorting between soluble and membrane-associated proteins. A membrane protein such as ALP may utilize a different set of sorting components than those used by the soluble vacuolar proteins. The

same may be true for dipeptidyl aminopeptidase B, another vacuolar membrane protein that has a topology similar to that of ALP (27). If different sorting components are used by membrane proteins, it may be possible to select mutants that are specifically defective in the sorting of this and possibly other membrane proteins. We are currently attempting to select such mutants by exploiting the Suc^- phenotype of a yeast strain harboring ALP-Inv hybrid proteins that are targeted to the vacuole. This approach has already proved useful for the selection of mutants defective in the sorting of soluble vacuolar hydrolases (10). Preliminary experiments show that we can select spontaneous Suc^+ mutants that missort the ALP-Inv191 hybrid protein to the cell surface.[2] Maturation of the wild-type ALP protein is also impaired in these mutants, consistent with missorting of the native hydrolase.

ALP is a useful model protein for examining the factors affecting interorganelle transfer and sorting of membrane proteins in eukaryotes. The proteinase A-dependent maturation of ALP (28) is unique for the vacuolar membrane proteins that have been characterized. This proteolytic cleavage serves as a convenient way to assess vacuolar delivery and is particularly useful in analyzing potential missorting mutants. In addition, the stability of CPY20/ALPΔ1-53 suggests that ALP should be amenable to a detailed mutational analysis of its *cis*-acting vacuolar sorting information. Saturation mutagenesis of this region may allow us to precisely define the relevant sequence and structural features of this new class of protein sorting signals.

Acknowledgments—We thank Lois Banta and Editte Gharakhanian for critically reading this manuscript.

REFERENCES

1. Kornfeld, S. (1986) *J. Clin. Invest.* **77,** 1–6
2. Davis, C. G., Lehrman, M. A., Russell, D. W., Anderson, R. G. W., Brown, M. S., and Goldstein, J. L. (1986) *Cell* **45,** 15–24
3. Santos, M. J., Imanaka, T., Shio, H., Small, G. M., and Lazarow, P. B. (1988) *Science* **239,** 1536–1538
4. Palade, G. (1975) *Science* **189,** 347–358
5. Kornfeld, S. (1987) *FASEB J.* **1,** 462–468
6. Schekman, R. (1985) *Annu. Rev. Cell Biol.* **1,** 115–143
7. Johnson, L. M., Bankaitis, V. A., and Emr, S. D. (1987) *Cell* **48,** 875–885
8. Valls, L. A., Hunter, C. P., Rothman, J. H., and Stevens, T. H. (1987) *Cell* **48,** 887–897
9. Klionsky, D. J., Banta, L. M., and Emr, S. D. (1988) *Mol. Cell. Biol.* **8,** 2105–2116
10. Robinson, J. S., Klionsky, D. J., Banta, L. M., and Emr, S. D. (1988) *Mol. Cell. Biol.* **8,** 4936–4948
11. Rothman, J. H., Howald, I., and Stevens, T. H. (1989) *EMBO J.* **8,** 2057–2065
12. Green, S. A., Zimmer, K.-P., Griffiths, G., and Mellman, I. (1987) *J. Cell Biol.* **105,** 1227–1240
13. Barriocanal, J. G., Bonifacino, J. S., Yuan, L., and Sandoval, I. V. (1986) *J. Biol. Chem.* **261,** 16755–16763
14. Waheed, A., Gottschalk, S., Hille, A., Krentler, C., Pohlmann, A., Braulke, T., Hauser, H., Geuze, H., and von Figura, K. (1988) *EMBO J.* **7,** 2351–2358
15. Klionsky, D. J., and Emr, S. D. (1989) *EMBO J.* **8,** 2241–2250
16. Hasselbeck, A., and Schekman, R. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 2017–2021
17. Bankaitis, V. A., Johnson, L. M., and Emr, S. D. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 9075–9079
18. Ballou, C. (1976) *Adv. Microb. Physiol.* **14,** 93–158
19. Trimble, R. B., Maley, F., and Chu, F. K. (1983) *J. Biol. Chem.* **258,** 2562–2567
20. Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54,** 631–664
21. Kaneko, Y., Hayashi, N., Toh-e, A., Banno, I., and Oshima, Y. (1987) *Gene (Amst.)* **58,** 137–148
22. von Heijne, G. (1986) *Nucleic Acids Res.* **14,** 4683–4690
23. Clark, D. W., Tkacz, J. S., and Lampen, J. O. (1982) *J. Bacteriol.* **152,** 865–873
24. Gething, M.-J., McCammon, K., and Sambrook, J. (1986) *Cell* **46,** 939–958
25. Bowman, E. J., Siebers, A., and Altendorf, K. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 7972–7976
26. Banta, L. M., Robinson, J. S., Klionsky, D. J., and Emr, S. D. (1988) *J. Cell Biol.* **107,** 1369–1383
27. Roberts, C. J., Pohlig, G., Rothman, J. H., and Stevens, T. H. (1989) *J. Cell Biol.* **108,** 1363–1373
28. Jones, E. W., Zubenko, G. S., and Parker, R. R. (1982) *Genetics* **102,** 665–667
29. Blachly-Dyson, E., and Stevens, T. H. (1987) *J. Cell Biol.* **104,** 1183–1191

[2] D. J. Klionsky and S. D. Emr, unpublished results.

Lysosome-associated protein transmembrane 4α (LAPTM4α) requires two tandemly arranged tyrosine-based signals for sorting to lysosomes

Douglas L. HOGUE*[1], Colin NASH†, Victor LING‡ and Tom C. HOBMAN†[2]

*Department of Biochemistry and Molecular Biology, Charles Tupper Building, Dalhousie University, Halifax, NS, Canada B3H 4H7, †Department of Cell Biology, Medical Sciences Building, University of Alberta, Edmonton, AB, Canada T6G 2H7, and ‡British Columbia Cancer Research Centre, 600 West 10th Avenue, Vancouver, BC, Canada V5Z 1L3

Lysosome-associated protein transmembrane 4α (LAPTM4α) and homologues comprise a family of conserved proteins, which are found in mammals, insects and nematodes. LAPTM4α functions to regulate the intracellular compartmentalization of amphipathic solutes and possibly the sensitivity of cells toward anthracyclines, antibiotics, ionophores, nucleobases and organic cations. This is similar to the multidrug-resistance phenotype exhibited by cells synthesizing high levels of P-glycoprotein. Accordingly, it is possible that LAPTM4α may be a suitable target for development of novel chemotherapeutic agents. LAPTM4α contains four putative membrane-spanning domains and a 55 amino acid C-terminal region that faces the cytoplasm. Localization of LAPTM4α to endosomes and lysosomes appears to be tightly controlled as transient high-level expression of LAPTM4α in cultured cells resulted in no detectable protein on the cell surface. Mutagenic analysis of the C-terminus of LAPTM4α indicated that two tandomly arranged tyrosine-containing motifs in the cytoplasmic domain are required for efficient localization of LAPTM4α to vesicles containing the lysosomal marker lysosomal glycoprotein 120. Although a number of membrane proteins that localize to endosomes/lysosomes contain more than one independently functioning sorting signal, to our knowledge, LAPTM4α is the first example of a membrane protein that requires two tandemly arranged tyrosine-based sorting signals for efficient localization in these compartments.

Key words: compartmentalization, late endosome, membrane protein.

INTRODUCTION

Lysosome-associated protein transmembrane 4α (LAPTM4α) is a four transmembrane-spanning protein that is relatively abundant in lysosome-enriched membranes isolated from rat liver, but whose endogenous levels of expression in various cultured cell lines are below the detection limits of immunofluorescence microscopy [1,2]. We recently demonstrated [1] LAPTM4α to be a resident protein of lysosomes and late endosomes by using indirect immunofluorescence microscopic analysis of transiently expressed wild-type LAPTM4α and N-terminal haemagglutinin epitope-tagged LAPTM4α in cultured cells. Functional studies employing the heterologous expression of LAPTM4α within the prevacuolar compartments of a drug-sensitive strain of *Saccharomyces cerevisiae* have demonstrated its ability to regulate cell sensitivity towards anthracyclines, antibiotics, dihydropyridines, ionophores, cytotoxic nucleobases and organic cations [2]. In addition, LAPTM4α was able to alter intracellular steroid distribution in this yeast strain as well as in cultured mammalian cells (D. Hogue, unpublished work). The multidrug-resistant phenotype is similar to that exhibited by mammalian cells which overexpress P-glycoprotein, an ATP-binding cassette membrane protein that mediates the cellular efflux of xenobiotics and steroids [3]. Together, these studies indicate that LAPTM4α functions to regulate the compartmentalization of amphiphilic solutes within lysosomes and late endosomes. Accordingly, it is possible that LAPTM4α may be a suitable target for development of novel chemotherapeutic agents.

LAPTM4α is one of the few type III membrane proteins currently identified to be a resident of lysosomes and late endosomes, the others being lysosomal glycoprotein (lgp)85/lysosomal integral membrane protein (LIMP) II, CD63/LIMP I, natural resistance-associated macrophage protein 2 (NRAMP2)/divalent metal transporter 1 ('DMT1') and cystinosin [4–7]. The functions of lgp85/LIMP II and CD63/LIMP I are currently unknown, but NRAMP2 is thought to transport free bivalent metal ions from the lumen of late endosomes and lysosomes into the cytosol [8]. Cystinosin is a seven-transmembrane domain glycoprotein, which is predicted to transport cystine from the lysosome lumen to the cytosol.

Considerable effort has been dedicated to the study of sorting signals responsible for localization of type I membrane proteins to lysosomes and late endosomes [9–13]. For example, single tyrosine-based and/or di-leucine motifs present within the short (10–30 amino acid) cytoplasmic tails of lgp120/lysosome-associated membrane protein ('LAMP I') [14], tyrosinase [15,16] and endolyn [17] serve as sorting signals. The spacing of these tyrosine-based motifs relative to the transmembrane-spanning helix is important for optimal sorting of type I membrane proteins to lysosomes [13,18], as is the nature of the Φ residue within the tyrosine-based motif (YXXΦ, where X is any residue and Φ is a large hydrophobic residue) [19]. These features, as well as the presence of multiple potential sorting motifs on a single protein, complicate the analysis of these motifs with regard to their specific roles in sorting to lysosomes and late endosomes. Studies of sorting signals within type III membrane proteins that are responsible for their lysosomal/late-endosomal localization are restricted to those performed on lgp85/LIMP II and cystinosin. The cytoplasmically exposed C-terminus of lgp85/LIMP II contains a di-leucine-based motif necessary for its

Abbreviations used: AP, adaptor protein complex; CHX, cycloheximide; EST, expressed sequence tag; LAPTM4α, lysosome-associated protein transmembrane 4α; lgp, lysosomal glycoprotein; LIMP, lysosomal integral membrane protein; VL motif, valine–leucine motif; wt, wild type.

[1] Deceased.

[2] To whom correspondence should be addressed (e-mail tom.hobman@ualberta.ca).

proper localization [20] and a di-acidic amino acid motif, which contributes towards its steady state-distribution between lysosomes, late endosomes and early endosomes [21]. Localization of cystinosin to lysosomes requires the presence of two tyrosine-based motifs, which are present on different cytoplasmically exposed regions of the protein [7]. Presently, nothing is known about how LAPTM4α is targeted to endosomes/lysosomes.

In the present study, our goal was to identify regions within LAPTM4α that serve as sorting signals for its trafficking to lysosomes and late endosomes. Two tandemly arranged tyrosine-based motifs within the C-terminus of LAPTM4α have been identified as being essential for proper lysosomal/late endosomal localization. Identification of additional mammalian, insect and nematode members of the LAPTM family has also allowed for a more detailed examination of these sorting signals and other features within the C-termini of these proteins.

MATERIALS AND METHODS

Generation of LAPTM4α expression vectors

The open reading frame of murine LAPTM4α (GenBank® accession number NM_008640) was PCR amplified (primer A, 5′-GCA*GGATC*CATGGTGTCCATGACTTTC-3′, where the *Bam*HI site is italicized, *Nco*I site is underlined and the translation initiation codon is in bold; and primer B, 5′-GCAGGTACCT-CAGGCAGGCAGGTAAGGA-3′, where the *Kpn*I site is underlined) and cloned into the *Bam*HI and *Kpn*I sites of pBlueScript II. Expand polymerase (Roche Molecular Biochemicals, Laval, QC, Canada) was used for all PCR reactions. Two complementary oligonucleotides (5′-CATGGAGGAGCAAA-AGCTTATTTCTGAAGAGGACTTGCT-3′ and 5′-CATG-AGCAAGTCCTCTCCAGAAATAAGCTTTTGCTCCTC-3′) encoding the c-myc epitope (MEEQKLISEEDLL, where single-letter amino-acid notation has been used and the epitope is underlined) were ligated into the unique *Nco*I site. The c-myc–LAPTM4α open reading frame was cloned as a *Bam*HI–*Kpn*I fragment into pcDNA3 to obtain pcDNA3–myc–LAPTM4α. An identical approach was employed to generate pcDNA3–myc–LAPTM4αΔCT, except that primer C (5′-GCAGGTACC-TCAAGCCTTAAAAATGATGAATAC-3′, where the translation stop codon is in bold and the *Kpn*I site is underlined) was substituted for primer B.

Generation of CD8–C-terminus (CD8–CT) expression vectors

A combination of PCR and overlap extension was used to fuse the coding regions for the extracellular and transmembrane domains of CD8 to the cytoplasmic domain of LAPTM4α. Briefly, the LAPTM4α cDNA was used as template in a PCR reaction with the following primers: forward primer 1, 5′-CCCTTTACTGCAACCACCTAATCAATTGTGTTTGG-AAC-3′; and reverse primer 2, 5′-ACC*GTCGAC*TCAGGCAG-GCAGGTAAGG-3′. The underlined bases in forward primer 1 correspond to the 3′ region of the CD8 transmembrane domain. Reverse primer 2 contains a *Sal*I site (bold and italic). The resulting 194 bp PCR product was purified and used as the reverse primer in a second PCR reaction using a human CD8 cDNA [22] as the template. The forward primer for the second PCR was a vector-specific oligonucleotide (5′-TACGGTGGGA-GGTCTATATAGC-3′). The second PCR product CD8–CT_{wt} (wild type), which encodes the CD8 extracellular and transmembrane domains of CD8 fused to the cytoplasmic domain of LAPTM4α, was digested with *Eco*RI and *Sal*I and then ligated into the mammalian expression vector pCMV5 [23]. The authenticity of the CD8–CT cDNA was confirmed by DNA sequencing.

Site-directed mutagenesis of the CD8–CT_{wt} expression vector and pcDNA3–myc–LAPTM4α was performed using the QuikChange Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA, U.S.A.) and pairs of complementary mutagenic oligonucleotides. DNA sequences of myc–LAPTM4α, myc–LAPTM4αΔCT, CD8–CT_{wt} and identity of all nucleotide mutations were confirmed by DNA sequencing.

Identification of LAPTM4β, *Bombyx mori* LAPTM, *Caenorhabditis elegans* LAPTM and *Drosophila melanogaster* LAPTM cDNAs

The amino acid sequence of LAPTM4α was compared with expressed sequence tag (EST) nucleotide databases from which a human EST cDNA (GenBank® accession number AA995835) and murine EST cDNA (GenBank® accession number AA168317) were identified, obtained from a commercial source and subjected to DNA sequencing. Anchor PCR amplification was performed using three different human cDNA libraries as independent templates to confirm the identity of the 5′ terminus of human LAPTM4β. The full-length cDNA sequences of human and murine LAPTM4β have been assigned the GenBank® accession numbers AF317417 and AF317418 respectively.

Searching of non-mammalian EST nucleotide databases was performed using regions of the mammalian LAPTM protein sequences. The *B. mori* cDNA clone e40437 (GenBank® accession number AU000351) was provided by Dr K. Mita (Genome Research Group, Tokyo, Japan), the *Caenorhabditis elegans* cDNA clone yk643e7 (GenBank® accession number AV195585) was provided by Dr Y. Kohara (National Institute of Genetics, Mishima, Japan), and the *D. melanogaster* cDNA clone GH27451 (GenBank® accession number AI516866) was obtained from a commercial source (ResGen, Burlington, ON, Canada). Clones were subjected to DNA sequencing and all were found to encode full-length proteins. The cDNA sequences of the LAPTM orthologues from *B. mori*, *D. melanogaster* and *C. elegans* have been assigned GenBank® accession numbers AF317420, AF317419 and AF317421 respectively. It should be noted that the coding regions of the *D. melanogaster* and *C. elegans* LAPTM homologues are incorrectly predicted and annotated in their respective genomic databases.

Cell culture

COS-1 cells obtained from the American Type Culture Collection (Bethesda, MD, U.S.A.) were grown in Dulbecco's modified Eagle's medium containing 10% (v/v) foetal bovine serum and plated at 1.3×10^5 cells/well in 35-mm wells containing 12-mm glass coverslips. Transfections were performed 24 h later by co-incubation with 2 μg of expression vector and 5 μl of FuGene 6 (Roche Molecular Biochemicals). At 24–40 h post-transfection, cells were processed for indirect immunofluorescence staining. Where indicated, protein synthesis was inhibited by the incubation of cells in 500 μM cycloheximide (CHX) for 3 h prior to processing for staining.

Indirect immunofluorescence microscopy

Cells were fixed with methanol at −20 °C and then processed for indirect immunofluorescence using mouse anti-myc (9E10),

mouse anti-CD8 (OKT8) and rabbit anti-(lgp120) antibodies, as described previously [24]. Cells were examined using a Zeiss Axioskope2 microscope or, where indicated, a Zeiss 510 confocal microscope. Multiple optical sections (0.5 μm) were collected from each sample and representative optical sections (X–Y orientation) are shown in the Figures.

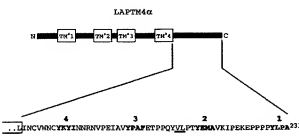


Figure 1 A schematic representation of murine LAPTM4α

The full-length protein is depicted at the top, with orientation of the predicted four transmembrane-spanning helices (TM) relative to the N- and C-termini. The amino acid sequence of the 55 amino acids comprising the hydrophilic carboxyl terminus of LAPTM4α is shown at the bottom. Potential tyrosine-based sorting motifs (bold) are indicated and numbered inwards from the C-terminus. The VL motif is underlined.

RESULTS

The C-terminus of LAPTM4α is necessary for its sorting to lysosomes and late endosomes

The hydrophilic C-terminus of LAPTM4α contains four tyrosine-based motifs (YXXΦ) and a single valine–leucine (VL) motif that could potentially function as sorting signals (Figure 1); therefore the role of this entire region in sorting was investigated. We employed LAPTM4α constructs that contained a myc-epitope tag at its N-terminus in these experiments, a location which has been shown previously [1] to tolerate the presence of a small epitope tag without altering the intracellular localization of the protein. Accordingly, the myc-tagged LAPTM4α distributed to a population of intracellular vesicles many of which were positive for lgp120, a resident membrane protein of lysosomes and late endosomes (Figures 2A–2C). Transfected cells were treated with CHX to reduce the pool of newly synthesized LAPTM4α in the perinuclear region, thus allowing us to examine more clearly the compartments in which this protein stably resides. CHX treatment did not alter the co-localization of LAPTM4α and lgp120 in vesicles (Figures 2D–2F), suggesting that LAPTM4α is a stable resident protein of lysosomes/late endosomes.

The deletion of the entire hydrophilic C-terminus from LAPTM4α led to a near-complete loss of its co-localization with

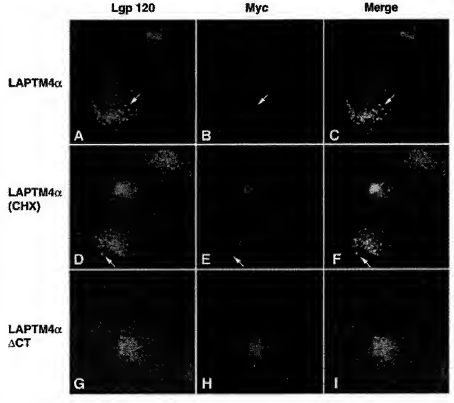


Figure 2 C-terminus of LAPTM4α is essential for its sorting to lysosomes and late endosomes

COS cells were transfected with expression vectors encoding myc-tagged LAPTM4α (**LAPTM4α**) (**A–F**) or myc-tagged LAPTM4α lacking the C-terminus (LAPTM4αΔCT) (**G–I**). (**D–F**) CHX (500 μM) treatment was performed prior to cell fixation and staining with rabbit anti-(lgp120) and mouse anti-myc antibodies. Primary antibodies were visualized with FITC-conjugated anti-(rabbit IgG) (**A**, **D** and **G**) and Texas Red-conjugated anti-(mouse IgG) (**B**, **E** and **H**) antibodies. (**C**, **F** and **I**) Merged confocal images, where co-localization of lgp120 and myc-tagged LAPTM4α is observed as yellow signals. Arrows denote vesicles that contain both lgp120 and LAPTM4α. Magnification, × 85.

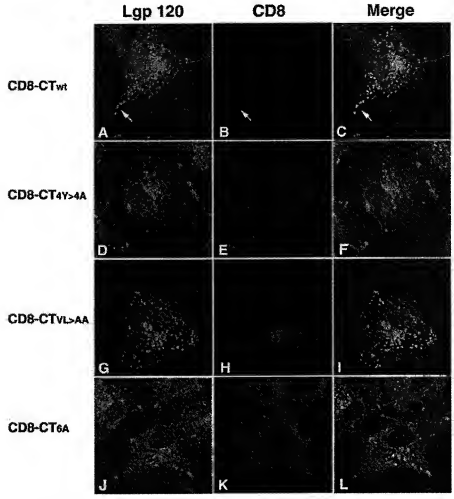


Figure 3 C-terminus of LAPTM4α contains tyrosine-based signals necessary for sorting to lysosomes and late endosomes

COS cells were transfected with expression vectors encoding CD8–CT_{wt} (**A–C**), CD8–$CT_{4Y>4A}$ (**D–F**), CD8–$CT_{VL>AA}$ (**G–I**) or CD8–CT_{6A} (**J–L**). Cells were treated with CHX (500 μM) for 3 h, fixed, and then stained with rabbit anti-(lgp120) and mouse anti-CD8 antibodies. Primary antibodies were visualized with FITC-conjugated anti-(rabbit IgG) (**A**, **D**, **G** and **J**) and Texas Red-conjugated anti-(mouse IgG) (**B**, **E**, **H** and **K**) antibodies. (**C**, **F**, **I** and **L**) Merged confocal images, where co-localization of lgp120 and CD8 fusion proteins is observed as yellow signals. Arrows denote vesicles that contain both lgp120 and CD8. Magnification, ×65.

lgp120 (Figures 2G–2I). The truncated protein, designated LAPTM4αΔCT, localized to juxtanuclear regions of transfected cells as well as the endoplasmic reticulum. Subsequent experiments revealed that the juxtanuclear pattern of LAPTM4αΔCT overlapped significantly with the resident Golgi membrane protein giantin (results not shown). This indicated that the hydrophilic C-terminus is necessary for the proper sorting of LAPTM4α to endocytic compartments.

Identification of individual sorting signals in the C-terminus of LAPTM4α

The results presented above indicated that the C-terminus of LAPTM4α was required for sorting to lysosomes, but did not reveal whether it was sufficient for sorting to these compartments. Moreover, the fact that LAPTM4αΔCT was arrested in a juxtanuclear compartment, which is proximal to lysosomes, raises the concern that the mutant protein is either misfolded/improperly assembled or that the C-terminal domain of LAPTM4α is required for efficient transport beyond the juxtanuclear region. Accordingly, we elected to address whether the C-terminal domain of LAPTM4α was able to re-direct the localization of a plasma-membrane protein to lysosomes. The type I plasma membrane protein CD8 has been used by us [25] and others [26] to analyse intracellular sorting signals. We have demonstrated previously [25] that wild-type CD8 protein expressed from the vector pCMV5 in cultured cells localizes to the plasma membrane. A chimaeric construct consisting of the luminal and transmembrane domains of CD8 fused to the C-terminal 55 amino acid residues of wild-type LAPTM4α (CD8–CT_{wt}) was expressed in transfected cells and examined by indirect immunofluorescence. The distribution of CD8–CT_{wt} was primarily restricted to a vesicle population, many of which were positive for lgp120 (Figures 3A–3C), thereby confirming that the

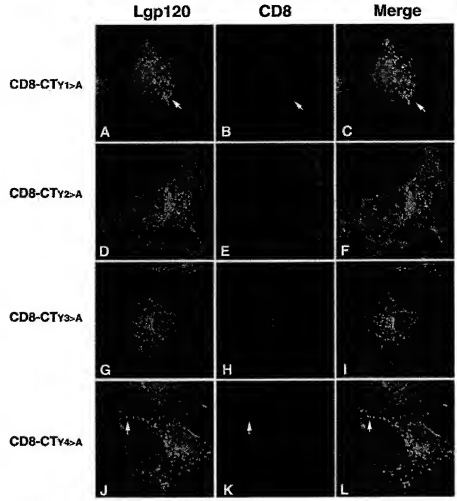


Figure 4 Mutation of tyrosine motifs 2 or 3 in the C-terminus of LAPTM4α leads to a partial loss of lysosomal localization

COS cells were transfected with expression vectors encoding CD8–$CT_{Y1>A}$ (**A–C**), CD8–$CT_{Y2>A}$ (**D–F**), CD8–$CT_{Y3>A}$ (**G–I**) or CD8–$CT_{4Y>4A}$ (**J–L**). Cells were treated with CHX (500 μM) for 3 h, fixed, and then stained with rabbit anti-(lgp120) and mouse anti-CD8 antibodies. Primary antibodies were visualized with FITC-conjugated anti-(rabbit IgG) (**A**, **D**, **G** and **J**) and Texas Red-conjugated anti-(mouse IgG) (**B**, **E**, **H** and **K**) antibodies. (**C**, **F**, **I** and **L**) Merged confocal images, where co-localization of lgp120 and CD8 fusion proteins is observed as yellow signals. Arrows denote vesicles that contain both lgp120 and CD8. Magnification, × 65.

C-terminus of LAPTM4α is sufficient for sorting to lysosomes/late endosomes.

The next objective was to identify amino acid motifs within the C-terminus of LAPTM4α that are responsible for its sorting to lysosomes. To study the potential sorting roles of the tyrosine-based motifs in the C-terminus of LAPTM4α (see Figure 1), all four tyrosine residues were mutated to alanine residues (CD8–$CT_{4Y>4A}$). Alanine substitution is often utilized in mutagenesis studies, because its small neutral side group of minimizes the introduction of spatial or electrostatic constraints. CD8–$CT_{4Y>4A}$ was localized primarily to the plasma membrane and exhibited very little overlap with lgp120-positive vesicles (Figures 3D–3F), revealing that one or more of the four tyrosine-based motifs was necessary for sorting to lysosomes. This prediction was supported by the observation that myc-tagged LAPTM4α containing alanine substitutions at all four tyrosine residues also failed to co-localize with lgp120-positive vesicles, but instead exhibited a juxtanuclear localization that partially overlapped with the Golgi-resident protein giantin (results not shown).

Di-leucine and leucine-based motifs also constitute important signals for the sorting of integral membrane proteins to lysosomes and late endosomes [27]. Although the C-terminus of LAPTM4α does not contain a di-leucine motif, the VL motif present may serve as a variant of the di-leucine motif, and for this reason we decided to determine if it functioned in sorting to endosomes/lysosomes (Figure 1). CD8–$CT_{VL>AA}$ was generated by alanine substitution of this VL motif, and its sub-cellular distribution was subsequently assessed by indirect immunofluorescence. Interestingly, CD8–$CT_{VL>AA}$ was distributed within lgp120-positive vesicles and the juxtanuclear region in addition to the plasma membrane (Figures 3G–3I). When alanine substitution of the four tyrosine residues and the VL motif was performed, the resulting protein (CD8–CT_{6A}) exhibited a pattern of localization very similar to that exhibited by CD8–$CT_{4Y>4A}$ (see Figures 3D–3F). The majority of CD8–CT_{6A} was distributed to the

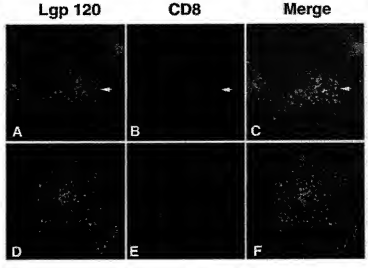


Figure 5 Mutation of tyrosine motifs 2 and 3 in the C-terminus of LAPTM4α leads to complete loss of lysosomal localization

COS cells were transfected with expression vectors encoding CD8–$CT_{Y1,4>A}$ (**A–C**) and CD8–$CT_{Y2,3>A}$ (**D–F**). Cells were treated with CHX (500 μM) for 3 h, fixed, and then stained with rabbit anti-(lgp120) and mouse anti-CD8 antibodies. Primary antibodies were visualized with FITC-conjugated anti-(rabbit IgG) (**A** and **D**) and Texas Red-conjugated anti-(mouse IgG) (**B** and **E**) antibodies. (**C** and **F**) Merged confocal images, where co-localization of lgp120 and CD8 fusion proteins is observed as yellow signals. Arrows denote vesicles that contain both lgp120 and CD8. Magnification, × 65.

plasma membrane, with very little being observed in lgp120-positive vesicles or the juxtanuclear region (Figures 3J–3L). This suggested that the VL motif may either play a role directly in LAPTM4α trafficking to endosomes/lysosomes or that mutagenesis of this motif perturbed the ability of adjacent tyrosine-based motifs to mediate proper sorting to these compartments. Indeed, alteration of amino acids adjacent to tyrosine-based sorting motifs can affect trafficking of membrane proteins possibly by regulating exposure of the signal to their cognate adaptor complexes [28,29].

The contribution of each tyrosine-based motif towards the sorting of CD8–CT_{wt} into lysosomes and late endosomes was examined next. Alanine substitutions were independently performed on the tyrosine residues within each of the four motifs, which are numbered inwards from the C-terminus of LAPTM4α (see Figure 1). CD8–$CT_{Y1>A}$, which harbours a substitution in motif 1, was localized to a lgp120-containing vesicle population that was indistinguishable from that exhibited by CD8–CT_{wt} (Figures 4A–4C). In contrast, substitutions of tyrosine residues in motifs 2 (CD8–$CT_{Y2>A}$) or 3 (CD8–$CT_{Y3>A}$) resulted in a significant loss of association with vesicular structures and consequent localization to the plasma membrane (Figures 4D–4I). Substitution in motif 4 (CD8–$CT_{Y4>A}$) did not significantly alter the transport of the chimaeric protein to lgp120-positive vesicles (Figures 4J–4L).

Tyrosine-based motifs 2 and 3 within CD8–CT_{wt} each appeared to significantly contribute towards its localization to lysosomes and late endosomes, whereas tyrosine-based motifs 1 and 4 did not. To assess any co-operative function of these two pairs of motifs, simultaneous alanine substitutions were performed. CD8–$CT_{Y1,4>A}$, which contained substitutions in both motif 1 and 4, localized to vesicles that contained lgp120 (Figures 5A–5C). This pattern of localization was indistinguishable from that of CD8–CT_{wt}, thereby confirming that motifs 1 and 4 do not appear to fulfil any sorting roles. Conversely, simultaneous substitutions in motifs 2 and 3 led to an apparent complete loss of CD8–$CT_{Y2,3>A}$ from lgp120-positive vesicles and re-distribution to the plasma membrane (Figures 5D–5F). These results suggest that in the context of the CD8 chimaeras, tyrosine motifs 2 and 3 of the LAPTM4α C-terminal domain act co-operatively to maintain the proteins in lysosomes and late endosomes, since substitution of either motif alone did not totally eliminate localization to lgp120-positive vesicles (Figures 4D–4I).

We also analysed the effect of single tyrosine to alanine substitutions on lysosomal targeting in the context of the myc–LAPTM4α protein (Figure 6). The results from these experiments support the conclusion that motifs 1 and 4 do not play significant roles in targeting of LAPTM4α to lysosomes (Figures 6A–6C and 6J–6L). Conversely, mutation of the tyrosine residues in motifs 2 or 3 resulted in accumulation of myc–LAPTM4α in the perinuclear region and concomitant loss of co-localization with lgp120-containing vesicles. Interestingly, unlike in the context of the CD8 chimaeras containing the LAPTM4α C-terminal domain, mutation of the tyrosines in either motif 2 or 3 resulted in complete loss of lysosomal-associated LAPTM4α (Figures 6D–6I). However, the results still support the previous data indicating that both motifs are required for efficient localization to lysosomes. In the context of LAPTM4α though, it appears that motifs 2 and 3 are also required to mediate efficient anterograde transport at the level of the endoplasmic reticulum and/or Golgi complex. Similarly, a recent study [30] showed that a YXXΦ motif, which was previously known to function in basolateral sorting of a viral membrane protein, is also required for efficient anterograde transport at the level of the endoplasmic reticulum and/or Golgi complex.

Conservation of sorting signals between members of the LAPTM family

The characterization of the tyrosine-based sorting motifs present within the C-terminus of LAPTM4α has been

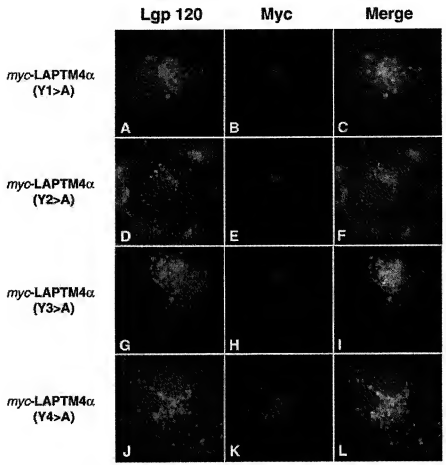


Figure 6 Tyrosine motifs 1 and 4 are not required for lysosomal targeting of LAPTM4α

COS cells were transfected with expression vectors encoding LAPTM4$\alpha_{Y1>A}$ (**A–C**), LAPTM4$\alpha_{Y2>A}$ (**D–F**), LAPTM4$\alpha_{Y3>A}$ (**G–I**) or LAPTM4$\alpha_{Y4>A}$ (**J–L**). Cells were treated with CHX (500 μM) for 3 h, fixed, and then stained with rabbit anti-(lgp120) and mouse anti-myc antibodies. Primary antibodies were visualized with FITC-conjugated anti-(rabbit IgG) (**A**, **D**, **G** and **J**) and Texas Red-conjugated anti-(mouse IgG) (**B**, **E**, **H** and **K**). (**C**, **F**, **I** and **L**) Merged images, where co-localization of lgp120 and LAPTM4α proteins is observed as yellow signals. Note the images were acquired using a Zeiss Axioskope2 microscope equipped with a SPOT™ digital camera. Magnification, ×65.

complemented by our identification and cloning of mammalian LAPTM4β as well as LAPTM orthologues from *B. mori* (silkworm), *C. elegans* (nematode) and *D. melanogaster* (fruitfly). Examination of completed genomic databases of *D. melanogaster* and *C. elegans* revealed that only a single LAPTM homologue is encoded within each of these species genomes. The mammalian LAPTM family currently consists of LAPTM4α, LAPTM4β and LAPTM5, which are classified according to their number of predicted transmembrane-spanning helices [1,31]. These three proteins exhibit 21–43% overall identity and each protein is highly conserved between mammalian species [1]. The insect homologues exhibit clusters of amino acid sequence conserved with the mammalian LAPTM proteins, whereas the *C. elegans* homologue has undergone significantly greater evolutionary divergence.

A comparison between the C-termini of these mammalian and non-mammalian LAPTM orthologues revealed that three of the four tyrosine-based motifs are conserved between all proteins, except for that of *C. elegans*. Disregarding the divergent *C. elegans* homologue, each of these motifs exists within a larger evolutionarily conserved domain of amino acid sequence (Figure 7). Tyrosine-based motifs 1 and 4 are not involved in sorting of LAPTM4α, but exist within the domains of PPPYXXA/V and VW/YXCYK/RXI/L respectively (where single-letter amino-acid notation has been used; the motifs are underlined and X represents any amino acid). These domains may be responsible for interacting with other binding proteins that regulate the function or stability of LAPTM4α. For example, we have shown ([32] and D. Hogue, unpublished work) that the 55 amino acid cytoplasmic domain of LAPTM4α binds to the ubiquitin ligase Nedd4. Although we have not mapped the regions of LAPTM4α that interact with Nedd4, the latter is known to interact with proline-rich tyrosine-containing motifs [33]. Of the two motifs important for the lysosomal/late-endosomal sorting of LAPTM4α, motif 2 (PXYE/DXA, where the motif is underlined and X represents any amino acid) is conserved between all LAPTM orthologues, and motif 3 (VYXXF/S, where the motif is underlined and X represents any amino acid) exists only in LAPTM4α and LAPTM4β. Alanine substitution of residues present at the −3 and −4 position (VL > AA) of motif 2 within LAPTM4α perturbed lysosomal sorting to a noticeable degree, indicating that this region may comprise an important

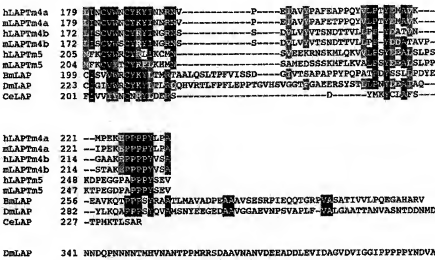


Figure 7 C-termini of LAPTM proteins contain conserved domains that encompass the tyrosine-based sorting motifs

A ClustalW alignment of the amino acid sequences of the hydrophilic C-termini of human and murine orthologues of LAPTM4α (GenBank® accession numbers NM_014713 and NM_008640 respectively), LAPTM4β (GenBank® accession numbers AF317417 and AF317418 respectively) and LAPTM5 (GenBank® accession numbers NM_006762 and NM_010686 respectively) as well as homologues from *B. mori* LAPTM (BmLAP, GenBank® accession number AF317420), *C. elegans* LAPTM (CeLAP (GenBank® accession number AF317421) and *D. melanogaster* LAPTM (DmLAP, GenBank® accession number AF317419) is shown. Shading indicates regions of identity (dark) or similarity (light). Consensus amino acid sequences of domains encompassing each of the four tyrosine-based motifs are shown.

feature of the sorting motif. However, the fact that tyrosine motif 2 is the most highly conserved among species, and mutation of this domain results in the most significant loss of targeting to lysosomes/endosomes, indicates that it may be the most important sorting determinant in the LAPTM family.

DISCUSSION

Membrane proteins destined to reside within, or traffic through, endosomes and lysosomes contain specific sorting signals within their hydrophilic cytoplasmic termini. Moreover, the presence of multiple sorting signals in the cytoplasmic domains of lysosome-associated integral membrane proteins is not altogether uncommon. Examples include lgp85/LIMP II, which contains di-leucine and di-acidic amino acid sorting motifs [20,21], tyrosinase, which utilizes di-leucine and tyrosine-based motifs for lysosomal targeting [15,16], cystinosin, which requires two tyrosine-based signals for efficient localization to lysosomes [7], and now LAPTM4α (the present study). The presence of multiple tyrosine-based sorting signals may be greatly beneficial for a protein whose cellular localization must be tightly restricted. For example, the tyrosine-based motifs may act as redundant sorting signals recognizable in multiple cellular compartments and thereby restrict the distribution of protein within the cell. In this respect, the two tandemly arranged tyrosine-based sorting signals present in LAPTM4α and LAPTM4β may function to tightly restrict the distribution of these proteins to lysosomes and late endosomes [1].

To our knowledge at least, LAPTM family members and cystinosin are the only membrane proteins known to contain multiple tyrosine-based motifs that co-function as signals for targeting to endosomes/lysosomes. A potentially significant difference between these two groups of proteins is that LAPTMs contain tandomly arranged motifs located within a contiguous 18 amino acid region (Figure 1), whereas the motifs in cystinosin are present on different cytoplasmic domains of the protein [7]. Accordingly, the LAPTM motifs may be part of a larger high-affinity binding site for proteins that ensures highly efficient sorting to endosomes and lysosomes at the level of the Golgi complex. It is tempting to speculate that the tyrosine-based motifs may also function to expedite anterograde transport of LAPTM4α at the level of the endoplasmic reticulum and/or Golgi complex by facilitating interactions with COPII complexes [34]. Indeed, evidence is emerging that, in addition to their well-recognized functions as post-Golgi sorting signals, tyrosine-based motifs play important roles in sorting at the level of the endoplasmic reticulum [30,35]. Since LAPTM4α regulates the compartmentalization of amphipathic molecules, its localization to endocytic membranes may be imperative, and mislocalization to other membrane compartments may be detrimental to cell physiology. For example, if LAPTM4α were mislocalized to the endoplasmic reticulum or the plasma membrane rather than endosomes/lysosomes, the associated transport activity of this protein could result in inappropriate transport of substrates or toxic metabolites into the endoplasmic reticulum or out of the cell, rather than into lysosomes where they would be degraded and the products recycled.

The conservation of the two tyrosine-based sorting signals between LAPTM4α and LAPTM4β, but not LAPTM5, suggests the cellular trafficking pattern of the latter protein may be subtly different. However, this may be significant in light of the fact that LAPTM5 expression is generally restricted to tissues and cells of haematopoetic origin [31], whereas LAPTM4α and LAPTM4β are expressed in many different types of tissues [36]. Thus the lack of one tyrosine-based sorting signal may enable LAPTM5 to fulfil a haematopoietic-specific role in lysosomes or a specialized

lysosome-derived compartment, such as secretory lysosomes [37]. Interestingly, none of the four YXXΦ motifs is preceded by a glycine residue, which is common feature of these motifs in lysosomal membrane proteins [38]. In addition, only one of the two required tyrosine-containing motifs (motif 3) conforms well to the classical YXXΦ, where Φ is a bulky hydrophobic amino acid, yet alteration of motif 2 had the greatest effect upon trafficking of LAPTM4α. This is similar to the atypical tyrosine motif YFPQA required for localization of cystinosin [7].

Tyrosine- and di-leucine-based motif recognition by adaptor complexes is mediated through their binding by different adaptor subunits, with tyrosine-based motifs being bound to the μ subunits and di-leucine motifs by the β subunits [39]. The association of a given adaptor complex with a tyrosine-based motif present in a membrane protein regulates the trafficking of the protein between different subcellular compartments. The adaptor protein complex (AP)-1 associates with clathrin-coated vesicles derived from the *trans*-Golgi network and endosomes, but its role in trafficking of lysosomal enzymes between endosomes and lysosomes has recently come into question [40]. AP-2 functions at the level of the plasma membrane to facilitate endocytosis of lysosomal enzymes and other proteins into clathrin-coated vesicles [39,41]. AP-3 is involved in the trafficking of certain lysosomal membrane proteins between the *trans*-Golgi network and lysosomes, but it is not clear if binding to the tyrosine-based sorting signals occurs at the level of the *trans*-Golgi network and/or endosomes [42–44]. Finally, AP-4 can interact with tyrosine-based motifs on certain lysosomal membrane proteins, but a direct role for this adaptor complex in transport of proteins to lysosomes has yet to be demonstrated [45]. Adaptor complexes, such as AP-1 and AP-3, also bind to the di-leucine-based sorting motifs of lysosomal/late-endosomal membrane proteins [11,44,46]. Members of the adaptor complex family exhibit different binding affinities for overlapping sets of tyrosine-based sorting motifs [47,48]. A recent study [49] indicates that amino acids that occupy positions +2 and +3 (relative to the tyrosine residue) are critical for differential binding to adaptor complexes; however, the authors suggest that the identities of amino acids at these positions cannot be used to generally predict which type of adaptor complex will bind to a given sorting signal. In the present study, however, we were unable to detect interactions between the μ1, μ2, μ3A or μ3B subunits with the hydrophilic C-terminus of LAPTM4α (results not shown).

In conclusion, the present study has demonstrated that the hydrophilic C-terminus of the type III membrane protein LAPTM4α is both necessary and sufficient for its proper localization to the lysosomal-endosomal compartment of mammalian cells. Deletion or mutagenesis of this C-terminal region induced an accumulation of LAPTM4α within the endoplasmic reticulum and Golgi, suggesting the C-terminus may also fulfil a role in facilitating the anterograde transport and/or complex assembly of LAPTM4α. Two tandemly arranged tyrosine-based motifs present in this C-terminal region have been identified as the major sorting signals responsible for the lysosomal localization of LAPTM4α. The fact that two interdependent sorting signals are present on LAPTM4α indicates that its localization to endosomes and lysosomes must be tightly regulated.

This work was supported by grants from the Canadian Institutes of Health Research (to T.C.H.) and the National Cancer Institute of Canada (to V.L.). T.C.H. is a Senior Medical Scholar supported by the Alberta Heritage Foundation for Medical Research. C.N. was supported by a summer studentship award from the Alberta Heritage Foundation for Medical Research. D.L.H. was supported by a British Columbia Health Research Foundation scholarship.

REFERENCES

1 Cabrita, M., Hobman, T., Hogue, D., King, K. and Cass, D. (1999) Mouse transporter protein (MTP), a membrane protein that regulates multidrug resistance, is localized to lysosomes. Cancer Res. **59**, 4899–4897
2 Hogue, D., Kerby, L. and Ling, V. (1999) A mammalian lysosomal membrane protein confers multidrug resistance upon expression in Saccharomyces cerevisiae. J. Biol. Chem. **274**, 12877–12882
3 Germann, U. A. (1996) P-glycoprotein – a mediator of multidrug resistance in tumour cells. Eur. J. Cancer **32A**, 927–944
4 Metzelaar, M. J., Wijngaard, P. L., Peters, P. J., Sixma, J. J., Nieuwenhuis, H. K. and Clevers, H. C. (1991) CD63 antigen. A novel lysosomal membrane glycoprotein, cloned by a screening procedure for intracellular antigens in eukaryotic cells. J. Biol. Chem. **266**, 3239–3245
5 Vega, M. A., Segui-Real, B., Garcia, J. A., Cales, C., Rodriguez, F., Vanderkerckhove, J. and Sandoval, I. V. (1991) Cloning, sequencing, and expression of a cDNA encoding rat LIMP II, a novel 74-kDa lysosomal membrane protein related to the surface adhesion protein CD36. J. Biol. Chem. **266**, 16818–16824
6 Vidal, S., Belouchi, A. M., Cellier, M., Beatty, B. and Gros, P. (1995) Cloning and characterization of a second human NRAMP gene on chromosome 12q13. Mamm. Genome **6**, 224–230
7 Cherqui, S., Kalatzis, V., Trugnan, G. and Antignac, C. (2001) The targeting of cystinosin to the lysosomal membrane requires a tyrosine-based signal and a novel sorting motif. J. Biol. Chem. **276**, 13314–13321
8 Tabuchi, M., Yoshimori, T., Yamaguchi, K., Yoshida, T. and Kishi, F. (2000) Human NRAMP2/DMT1, which mediates iron transport across endosomal membranes, is localized to late endosomes and lysosomes in HEp-2 cells. J. Biol. Chem. **275**, 22220–22228
9 Hunziker, W. and Geuze, H. J. (1996) Intracellular trafficking of lysosomal membrane proteins. BioEssays **18**, 379–389
10 Marks, M. S., Woodruff, L., Ohno, H. and Bonifacino, J. S. (1996) Protein targeting by tyrosine- and di-leucine-based signals: evidence for distinct saturable components. J. Cell Biol. **135**, 341–354
11 Honing, S., Sandoval, I. V. and von Figura, K. (1998) A di-leucine-based motif in the cytoplasmic tail of LIMP-II and tyrosinase mediates selective binding of AP-3. EMBO J. **17**, 1304–1314
12 Reaves, B. J., Banting, G. and Luzio, J. P. (1998) Luminal and transmembrane domains play a role in sorting type I membrane proteins on endocytic pathways. Mol. Biol. Cell **9**, 1107–1122
13 White, S., Hatton, S. R., Siddiqui, M. A., Parker, C. D., Trowbridge, I. S. and Collawn, J. F. (1998) Analysis of the structural requirements for lysosomal membrane targeting using transferrin receptor chimeras. J. Biol. Chem. **273**, 14355–14362
14 Guarnieri, F. G., Arterburn, L. M., Penno, M. B., Cha, Y. and August, J. T. (1993) The motif Tyr-X-X-hydrophobic residue mediates lysosomal membrane targeting of lysosome-associated membrane protein 1. J. Biol. Chem. **268**, 1941–1946
15 Calvo, P. A., Frank, D. W., Bieler, B. M., Berson, J. F. and Marks, M. S. (1999) A cytoplasmic sequence in human tyrosinase defines a second class of di-leucine-based sorting signals for late endosomal and lysosomal delivery. J. Biol. Chem. **274**, 12780–12789
16 Simmen, T., Schmidt, A., Hunziker, W. and Beermann, F. (1999) The tyrosinase tail mediates sorting to the lysosomal compartment in MDCK cells via a di-leucine and a tyrosine-based signal. J. Cell Sci. **112**, 45–53
17 Ihrke, G., Gray, S. R. and Luzio, J. P. (2000) Endolyn is a mucin-like type I membrane protein targeted to lysosomes by its cytoplasmic tail. Biochem. J. **345**, 287–296
18 Rohrer, J., Schweizer, A., Russell, D. and Kornfeld, S. (1996) The targeting of Lamp1 to lysosomes is dependent on the spacing of its cytoplasmic tail tyrosine sorting motif relative to the membrane. J. Cell Biol. **132**, 565–576
19 Gough, N. R., Zweifel, M. E., Martinez-Augustin, D., Aguilar, R. C., Bonifacino, J. S. and Fambrough, D. M. (1999) Utilization of the indirect lysosome targeting pathway by lysosome-associated membrane proteins (LAMPs) is influenced largely by the C-terminal residue of their GYXOphi targeting signals. J. Cell Sci. **112**, 4257–4269
20 Sandoval, I. V., Arredondo, J. J., Alcalde, J., Gonzalez Noriega, A., Vandekerckhove, J., Jimenez, M. A. and Rico, M. (1994) The residues Leu(IIe)475-IIe(Leu, Val, Ala)476, contained in the extended carboxyl cytoplasmic tail, are critical for targeting of the resident lysosomal membrane protein LIMP II to lysosomes. J. Biol. Chem. **269**, 6622–6631
21 Tabuchi, N., Akasaki, K. and Tsuji, H. (2000) Two acidic amino acid residues, Asp^{479} and Glu^{471}, contained in the carboxyl cytoplasmic tail of a major lysosomal membrane protein, LGP85/LIMP II, are important for its accumulation in secondary lysosomes. Biochem. Biophys. Res. Commun. **270**, 557–563
22 Liltman, D. R., Thomas, Y., Maddon, P. J., Chess, L. and Axel, R. (1985) The isolation and sequence of the gene encoding T8: a molecule defining functional classes of T lymphocytes. Cell (Cambridge, Mass.) **40**, 237–246

23 Andersson, S., Davis, D. L., Dahlback, H., Jornvall, H. and Russell, D. W. (1989) Cloning, structure, and expression of the mitochondrial cytochrome P-450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme. J. Biol. Chem. **264**, 8222–8229

24 Beatch, M. D. and Hobman, T. C. (2000) Rubella virus capsid associates with host cell protein p32 and localizes to mitochondria. J. Virol. **74**, 5569–5576

25 Hobman, T., Lemon, H. and Jewell, K. (1997) Characterization of an endoplasmic reticulum retention signal in the rubella virus E1 glycoprotein. J. Virol. **71**, 7670–7680

26 Ponnambalam, S., Rabouille, C., Luzio, J. P., Nilsson, T. and Warren, G. (1994) The TGN38 glycoprotein contains two non-overlapping signals that mediate localization to the *trans*-Golgi network. J. Cell Biol. **125**, 253–268

27 Geisler, C., Dietrich, J., Nielsen, B. L., Kastrup, J., Lauritsen, J. P., Odum, N. and Christensen, M. D. (1998) Leucine-based receptor sorting motifs are dependent on the spacing relative to the plasma membrane. J. Biol. Chem. **273**, 21316–21323

28 Lehmann, L. E., Eberle, W., Krull, S., Prill, V., Schmidt, B., Sander, C., von Figura, K. and Peters, C. (1992) The internalization signal in the cytoplasmic tail of lysosomal acid phosphatase consists of the hexapeptide PGYRNV. EMBO J. **11**, 4391–4399

29 Prill, V., Lehmann, L., von Figura, K. and Peters, C. (1993) The cytoplasmic tail of lysosomal acid phosphatase contains overlapping but distinct signals for basolateral sorting and rapid internalization in polarized MDCK cells. EMBO J. **12**, 2181–2193

30 Sevier, C. S., Weisz, O. A., Davis, M. and Machamer, C. E. (2000) Efficient export of the vesicular stomatitis virus G protein from the endoplasmic reticulum requires a signal in the cytoplasmic tail that includes both tyrosine-based and di-acidic motifs. Mol. Biol. Cell **11**, 13–22

31 Adra, C. N., Zhu, S., Ko, J. L., Guillemot, J. C., Cuervo, A. M., Kobayashi, H., Horiuchi, T., Lelias, J. M., Rowley, J. D. and Lim, B. (1996) LAPTM5: a novel lysosomal-associated multispanning membrane protein preferentially expressed in hematopoietic cells. Genomics **35**, 328–337

32 Rotin, D., Staub, D. and Haguenauer-Tsapis, R. (2000) Ubiquitination and endocytosis of plasma membrane proteins: role of Nedd4/Rsp5p family of ubiquitin-protein ligases. J. Membr. Biol. **176**, 1–17

33 Kasanov, J., Pirozzi, G., Uveges, A. J. and Kay, B. K. (2001) Characterizing Class I WW domains defines key specificity determinants and generates mutant domains with novel specificities. Chem. Biol. **8**, 231–241

34 Barlowe, C., Orci, L., Yeung, T., Hosobuchi, M., Hamamoto, S., Salama, N., Rexach, M. F., Ravazzola, M., Amherdt, M. and Schekman, R. (1994) COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. Cell (Cambridge, Mass.) **77**, 895–907

35 Nufer, D., Guldbrandsen, S., Degen, M., Kappeler, F., Paccaud, J. P., Tani, K. and Hauri, H. P. (2002) Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. J. Cell Sci. **115**, 619–628

36 Hogue, D., Ellison, M., Young, J. and Cass, C. (1996) Identification of a novel membrane transporter associated with intracellular membranes by phenotypic complementation in the yeast *Saccharomyces cerevisiae*. J. Biol. Chem. **271**, 9801–9808

37 Stinchcombe, J. C. and Griffiths, G. M. (1999) Regulated secretion from hemopoietic cells. J. Cell Biol. **147**, 1–6

38 Heilker, R., Spiess, M. and Crottet, P. (1999) Recognition of sorting signals by clathrin adaptors. BioEssays **21**, 558–567

39 Bonifacino, J. S. and Dell'Angelica, E. C. (1999) Molecular bases for the recognition of tyrosine-based sorting signals. J. Cell Biol. **145**, 923–926

40 Meyer, C., Zizioli, D., Lausmann, S., Eskelinen, E. L., Hamann, J., Saftig, P., von Figura, K. and Schu, P. (2000) μ1A-adaptin-deficient mice: lethality, loss of AP-1 binding and rerouting of mannose 6-phosphate receptors. EMBO J. **19**, 2193–2203

41 Rohn, W. M., Rouille, Y., Waguri, S. and Hollack, B. (2000) Bi-directional trafficking between the *trans*-Golgi network and the endosomal/lysosomal system. J. Cell Sci. **113**, 2093–2101

42 Dell'Angelica, E. C., Mullins, C. and Bonifacino, J. S. (1999) AP-4, a novel protein complex related to clathrin adaptors. J. Biol. Chem. **274**, 7278–7285

43 Hirst, J., Bright, N. A., Rous, B. and Robinson, M. S. (1999) Characterization of a fourth adaptor-related protein complex. Mol. Biol. Cell **10**, 2787–2802

44 Le Borgne, R. and Hoflack, B. (1998) Protein transport from the secretory to the endocytic pathway in mammalian cells. Biochim. Biophys. Acta **1404**, 195–209

45 Aguilar, R. C., Boehm, M., Gorshkova, I., Crouch, R. J., Tomita, K., Saito, T., Dhno, H. and Bonifacino, J. S. (2001) Signal-binding specificity of the μ4 subunit of the adaptor protein complex AP-4. J. Biol. Chem. **276**, 13145–13152

46 Fujita, H., Saeki, M., Yasunaga, K., Ueda, T., Imoto, T. and Himeno, M. (1999) *In vitro* binding study of adaptor protein complex (AP-1) to lysosomal targeting motif (LI-motif). Biochem. Biophys. Res. Commun. **255**, 54–58

47 Dhno, H., Fournier, M. C., Poy, G. and Bonifacino, J. S. (1996) Structural determinants of interaction of tyrosine-based sorting signals with the adaptor medium chains. J. Biol. Chem. **271**, 29009–29015

48 Dhno, H., Aguilar, R. C., Yeh, D., Taura, D., Saito, T. and Bonifacino, J. S. (1998) The medium subunits of adaptor complexes recognize distinct but overlapping sets of tyrosine-based sorting signals. J. Biol. Chem. **273**, 25915–25921

49 Obermuller, S., Klecke, C., von Figura, K. and Honing, S. (2002) The tyrosine motifs of Lamp1 and LAP determine their direct and indirect trafficking to lysosomes. J. Cell Sci. **115**, 185–194

Received 31 January 2002/13 March 2002; accepted 29 April 2002
Published as BJ Immediate Publication 29 April 2002, DOI 10.1042/BJ20020205

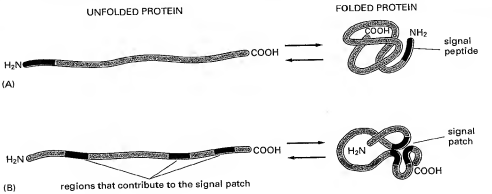


Figure 12–8 Two ways that a sorting signal can be built into a protein. (A) The signal resides in a single discrete stretch of amino acid sequence, called a *signal peptide*, that is exposed in the folded protein. Signal peptides often occur at the end of the polypeptide chain (as shown), but they can also be located elsewhere. (B) A *signal patch* can be formed by the juxtaposition of amino acids from regions that are usually physically separated before the protein folds (as shown); alternatively, separate patches on the surface of the folded protein that are spaced a fixed distance apart could form the signal. In either case the transport signal depends on the three-dimensional conformation of the protein, which makes it difficult to locate the signal precisely.

protein's surface that forms when the protein folds up. The amino acid residues that comprise this **signal patch** can be distant from one another in the linear amino acid sequence, and they generally remain in the finished protein (Figure 12–8). Signal peptides are used to direct proteins from the cytosol into the ER, mitochondria, chloroplasts, peroxisomes, and nucleus, and they are also used to retain soluble proteins in the ER. Signal patches identify certain enzymes that are to be marked with specific sugar residues that then direct them from the Golgi apparatus into lysosomes; signal patches are also used in other sorting steps that have been less well characterized.

Different types of signal peptides are used to specify different destinations in the cell. Proteins destined for initial transfer to the ER usually have a signal peptide at their amino terminus, which characteristically includes a sequence composed of about 5 to 10 hydrophobic amino acids. Most of these proteins will in turn pass from the ER to the Golgi apparatus, but those with a specific sequence of four amino acids at their carboxyl terminus are retained as permanent ER residents. Proteins destined for mitochondria have signal peptides of yet another type, in which positively charged amino acids alternate with hydrophobic ones. Proteins destined for peroxisomes usually have a specific signal sequence of three amino acids at their carboxyl terminus. Many proteins destined for the nucleus carry a signal peptide formed from a cluster of positively charged amino

Table 12–3 Some Typical Signal Peptides

Function of Signal Peptide	Example of Signal Peptide
Import into ER	$^{+}H_3N$-Met-Met-Ser-Phe-Val-Ser-Leu-Leu-Leu-Val Gly-Ile-Leu-Phe-Trp-Ala-Thr-Glu-Ala-Glu-Gln-Leu-Thr-Lys-Cys-Glu-Val-Phe-Gln-
Retain in lumen of ER	-Lys-Asp-Glu-Leu-COO^{-}
Import into mitochondria	$^{+}H_3N$-Met-Leu-Ser-Leu-Arg-Gln-Ser-Ile-Arg-Phe-Phe-Lys-Pro-Ala-Thr-Arg-Thr-Leu-Cys-Ser-Ser-Arg-Tyr-Leu-Leu-
Import into nucleus	-Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val-
Import into peroxisomes	-Ser-Lys-Leu-
Attach to membranes via the covalent linkage of a myristic acid to the amino terminus	$^{+}H_3N$-Gly-Ser-Ser-Lys-Ser-Lys-Pro-Lys-

Positively charged amino acids are shown in *red* and negatively charged amino acids in *green*. An extended block of hydrophobic amino acids is enclosed in a *yellow* box. H_3N^{+} indicates the amino terminus of a protein; COO^{-} indicates the carboxyl terminus.

Protein Targeting and Degradation

The eukaryotic cell is made up of many structures, compartments, and organelles, each with specific functions requiring distinct sets of proteins and enzymes. The synthesis of almost all these proteins begins on free ribosomes in the cytosol. How are these proteins directed to their final cellular destinations?

The answer to this question is at once complex, fascinating, and unfortunately incomplete. Enough is known, however, to outline many key steps in this process. Proteins destined for secretion, integration in the plasma membrane, or inclusion in lysosomes generally share the first few steps of a transport pathway that begins in the endoplasmic reticulum. Proteins destined for mitochondria, chloroplasts, or the nucleus each use separate mechanisms, and proteins destined for the cytosol simply remain where they are synthesized. The pathways by which proteins are sorted and transported to their proper cellular location are often referred to as **protein targeting** pathways.

Günter Blobel

The most important element in all of these targeting systems (with the exception of cytosolic and nuclear proteins) is a short amino acid sequence at the amino terminus of a newly synthesized polypeptide called the **signal sequence.** This signal sequence, whose function was first postulated by David Sabatini and Günter Blobel in 1970, directs a protein to its appropriate location in the cell and is removed during transport or when the protein reaches its final destination. In many cases, the targeting capacity of particular signal sequences has been confirmed by fusing the signal sequence from one protein, say protein A, to a different protein B, and showing that the signal directs protein B to the location where protein A is normally found.

The selective degradation of proteins no longer needed in the cell also relies largely on a set of molecular signals embedded in each protein's structure; most of these signals are not yet understood. The final part of this chapter is devoted to the processes of targeting and degradation, with emphasis on the underlying signals and molecular regulation that are so crucial to cellular metabolism. Except where noted, the focus is on eukaryotic cells.

Posttranslational Modification of Many Eukaryotic Proteins Begins in the Endoplasmic Reticulum

Perhaps the best-characterized targeting system begins in the endoplasmic reticulum (ER). Most lysosomal, membrane, or secreted proteins have an amino-terminal signal sequence that marks them for translocation into the lumen of the ER. More than 100 signal sequences for proteins in this group have been determined (Fig. 26–35). The se-

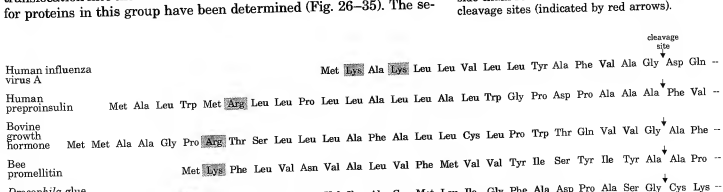
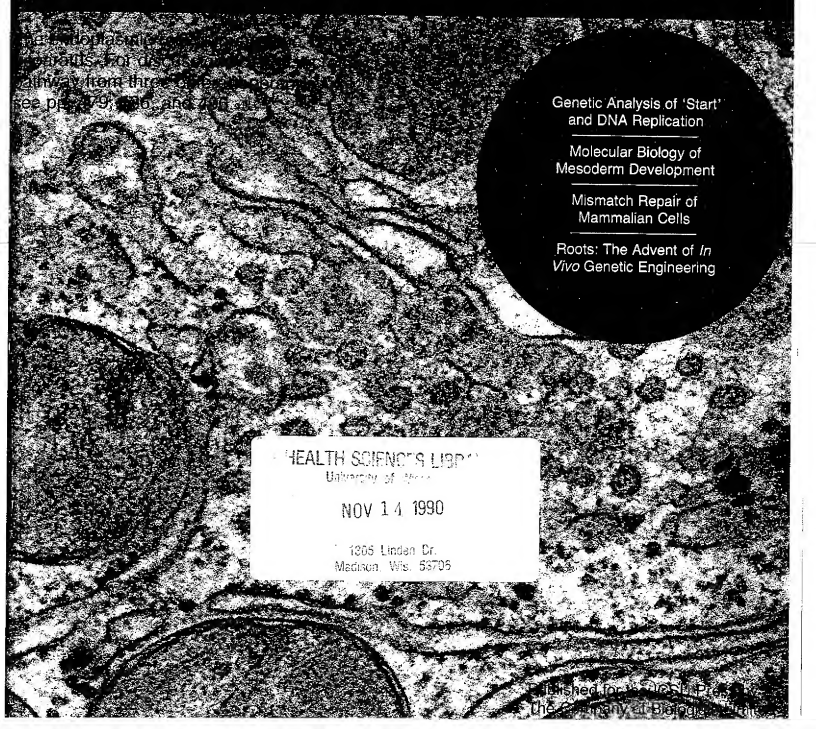


Figure 26–35 Amino-terminal signal sequences of some eukaryotic proteins, directing translocation into the endoplasmic reticulum. The hydrophobic core (yellow) is preceded by one or more basic residues (blue). Note the presence of polar and short-side-chain residues immediately preceding the cleavage sites (indicated by red arrows).

Volume 12 October 1990 No. 10

Advances in Molecular, Cellular and Developmental Biology

BioEssays



Genetic Analysis of 'Start' and DNA Replication

Molecular Biology of Mesoderm Development

Mismatch Repair of Mammalian Cells

Roots: The Advent of *In Vivo* Genetic Engineering

Published for BioEssays Readers by the Company of Biologists Limited

BioEssays is co-sponsored by the following members of the International Council of Scientific Unions:

International Union of Biological Sciences (IUBS)
International Union of Physiological Sciences (IUPS)
International Union of Pure and Applied Biophysics (IUPAB)
International Union of Pharmacology (IUPHAR)
International Union of Nutritional Sciences (IUNS)
International Union of Immunological Sciences (IUIS)
International Union of Microbiological Sciences (IUMS)
Committee on Genetic Experimentation (COGENE)
International Biosciences Network (IBN)

Contents

Cover Photograph

Region of interface between the endoplasmic reticulum and Golgi apparatus of rat liver where small vesicles bleb from transitional ER regions, migrate and fuse with the cis Golgi apparatus. For two different perspectives on ER to Golgi transport, see the articles by Newman and Ferro-Novick (pp. 485–491) and Locke (pp. 495–501); for a discussion of the molecular biology of protein transport into the ER, see the article by Nothwehr and Gordon (pp. 479–484).
(Photograph courtesy of Dr James Morre, Purdue University; reproduced with permission from Morre *et al.* (1989). *Protoplasma* 153, 1–13.)

Subscriptions

BioEssays (ISSN 0265-9247) is published by The Company of Biologists Limited, UK. One volume of 12 parts will be published in 1990. The subscription price includes air speeded delivery to most countries. Individuals placing personal subscriptions must certify that the journal is for their personal use. Institutional subscription prices: £110 (or £135 for airmail, direct); $199. Personal subscriptions: £44; $70.

Orders, which must be accompanied by payment or credit card authorization with details, may be sent to any bookseller, subscription agent or direct to The Company of Biologists Limited, Bidder Building, 140 Cowley Road, Cambridge, CB4 4DL, UK. US Postmaster: Send address corrections to: BioEssays, c/o Virgin Atlantic Cargo, Building 150, Newark International Airport, Newark, New Jersey, NJ 07114, USA.

Back volumes: Volumes 1–9 are available from Swets & Zeitlinger BV, P.O. Box 830, 2160 SZ Lisse/Holland. Volumes 10 and 11, which were published in 1989, are available from The Company of Biologists Limited at the address listed above.

Targeting of Proteins into the Eukaryotic Secretory Pathway: Signal Peptide Structure/Function Relationships

Steven F. Nothwehr and Jeffrey I. Gordon

Summary

Much progress has been made in recent years regarding the mechanisms of targeting of secretory proteins to, and across, the endoplasmic reticulum (ER) membrane. Many of the cellular components involved in mediating translocation across this bilayer have been identified and characterized. Polypeptide domains of secretory proteins, termed signal peptides, have been shown to be necessary, and in most cases sufficient, for entry of preproteins into the lumen of the ER. These NH_2-terminal segments appear to serve multiple roles in targeting and translocation. The structural features which mediate their multiple functions are currently the subject of intense study.

Introduction

Almost all eukaryotic protein synthesis occurs in the cytosol, yet many proteins exist in non-cytosolic locations such as the nucleus, mitochondria, chloroplast, peroxisomes, and the secretory pathway. The transport of proteins to these compartments involves targeting to, and translocation across, the 'appropriate' membrane bilayer. The 'cellular address code' which enables non-cytosolic proteins to be transported to their correct destination is contained within discrete topogenic signals[1,2]. These signals differ in their structure and location within the polypeptide, depending upon the final destination of the protein. Proteins which are destined to enter the secretory pathway *via* the ER were first shown by Milstein[3] and Blobel and Dobberstein[4,5] to be synthesized with NH_2-terminal extensions called signal peptides (or signal sequences). According to the signal hypothesis[4,5], the signal peptide serves as a sorting signal that targets the nascent secretory protein to sites of translocation on the ER membrane, where it is subsequently proteolytically removed from the mature chain. Over the last several years much evidence in favor of this hypothesis has accumulated[6]. Biochemical dissections of a mammalian *in vitro* translocation system, and more recently yeast genetic studies in *S. cerevisiae*, have allowed identification of several components involved in this process. Comparison of the primary sequences of naturally occurring signal peptides has revealed common features. The relationship of these features to function has been tested by assaying the activity of mutant signal peptides *in vivo* and *in vitro*. This review will focus primarily on structure/function relationships in eukaryotic signal peptides. However, the mechanism of translocation across the plasma membrane of prokaryotes and the role of prokaryotic signal sequences is presumed to be very similar to that of eukaryotes since each of the two systems are often able to recognize, translocate, and correctly process the same secretory proteins[7].

Structural Features of Signal Peptides

Most of what we know about signal peptide structure is derived from analysis of naturally occurring signal sequences[8,9]. Such comparisons indicated that signal peptides are quite variable in length (15 to as many as 50 amino acids) and contain very little identity in their primary structures. However, they do possess three general domains. The n, h, and c-regions have been defined on the basis of the physical-chemical characteristics of their constituent amino acids (Fig. 1). The NH_2-terminal n-region varies widely in length but typically contains amino acids which contribute a net positive charge to this domain. The h-region, or hydrophobic core, is made up of a block of 7–16 hydrophobic amino acids. This hydrophobic core is followed by 4–6 relatively polar amino acids that compose the COOH-terminal, or c-region. Often proline and glycine residues are found in the c-region, particularly at the h/c boundary. These residues may participate in β-turn structures. Small, neutral residues are typically encountered at the -3 and especially at the -1 positions (the numbering system is relative to the site of signal peptidase cleavage).

Since very little primary sequence conservation is apparent between signal peptides, it has been suggested that common secondary structural features are important for their function. Physical studies of isolated signal peptides indicate that they are capable of adopting multiple conformations depending on the environment (see ref. 6 for review). In general, these synthetic signal peptides exhibit little ordered structure in aqueous solvents as monitored by circular dichroism. In nonpolar solvents, they usually have significant amounts

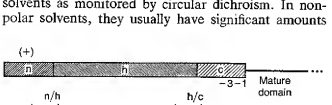


Fig. 1. Tripartite domain structure of eukaryotic signal peptides. Database analyses of the primary structures of naturally occurring signal peptides have identified three common domains. The n, h, and c-regions typically contain basic, hydrophobic, and polar residues, respectively. The boundary between the h and c-regions is often punctuated by Pro or Gly residues that are predicted to participate in β-turn structures. The -3 and -1 positions usually contain small, neutral residues.

of α-helical content although β-sheet structure has also been reported[10]. Nuclear magnetic resonance analyses of the bacterial LamB signal peptide indicate that, in membrane-like environments, it adopts α-helical structure, with the most stable helical regions located in the hydrophobic core domain[11].

Preprotein Interactions During Targeting, Translocation, and Cleavage

A schematic model of the currently envisioned pathway for co-translational targeting, translocation, and processing of eukaryotic preproteins is shown in Fig. 2. This model is based primarily on observations made from *in vitro* translation/translocation systems (for review see ref. 12).

The signal peptide of a nascent eukaryotic preprotein is initially recognized by the signal recognition particle (SRP) after chain elongation has produced a nascent preprotein of ~70 amino acids. This small cytoplasmic ribonucleoprotein contains a 300 nucleotide 7SL RNA and 6 polypeptide chains. A direct interaction between the signal peptide of preprolactin and the 54 kD subunit of SRP has been demonstrated by photochemical crosslinking studies. Interestingly, the 54 kD subunit contains a putative GTP binding region and a methionine-rich domain. The methionine-rich domain is predicted to contain a series of amphipathic helices having methionine residues located predominantly on one face. Walter and co-workers[13] therefore proposed that the relatively flexible methionine side chains may line a signal peptide binding pocket (in a fashion analogous to the bristles of a brush). The flexibility of such a binding site could account for the ability of SRP to bind an array of widely divergent signal peptides, although at present there is no direct evidence for this model. An initial, functional consequence of SRP binding to the signal peptide is a pausing or arrest of translation. This may increase the length of time that the precursor remains in an unfolded conformation compatible with translocation. A domain of SRP, separate from the signal recognition and translational arrest domains, binds to the docking protein (or SRP-receptor), an integral membrane protein of the ER membrane. In this way, polysomes engaged in the synthesis of proteins that contain signal peptides, are targeted to the ER membrane. Upon binding to the docking protein, the SRP-induced translational arrest is released and the nascent chain inserts into the membrane bilayer.

The insertion and subsequent translocation of the nascent polypeptide are poorly understood events. However, it is known that binding of GTP to the docking protein is involved in the process[14]. GTP hydrolysis has been hypothesized to act as a timing switch to ensure vectorial protein transport and/or proper fidelity of signal recognition[13,15]. The discovery that the 54 kD subunit of SRP also contains a putative GTP binding domain is consistent with the latter idea. Binding and hydrolysis of GTP is known to induce switching between two protein conformations[16] inviting speculation that GTP binding/hydrolysis by SRP and/or docking protein may regulate recognition and release of signal peptides.

Although translocation of most mammalian proteins appears to be tightly coupled to translation, recent evidence has shown that certain eukaryotic proteins (e.g. yeast prepro-α-factor) are capable of being translocated post-translationally[17]. For these proteins, a lack of stable tertiary structure is thought to be critical

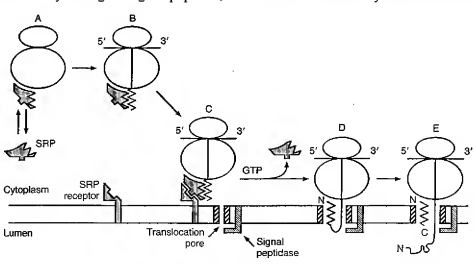


Fig. 2. Model of the early events of eukaryotic protein secretion. The affinity of SRP for ribosomes (*A*) is dramatically increased upon translation of mRNA encoding a secretory protein due to direct binding of SRP with the nascent protein's signal peptide (zigzag line). Binding of SRP to the signal sequence results in translational arrest (*B*). Interaction of SRP with SRP-receptor (docking protein) targets the SRP/nascent chain/ribosome complex to the ER membrane (*C*). In a reaction that requires GTP, SRP is released from the ribosome-bound nascent chain, translation resumes, and the nascent polypeptide is inserted into a hypothesized translocation pore in the membrane bilayer (*D*). The signal peptide cleavage site is recognized by signal peptidase located on the luminal side of the membrane. The remaining mature polypeptide domain is then translocated co-translationally across the ER membrane (E).

for maintaining (prolonging) their 'translocation competent' state. Cytoplasmic factors have been identified in both prokaryotes (trigger factor and SecB) and in lower eukaryotes (hsp 70) that bind to fully synthesized proteins and promote their translocation. These proteins, known as 'chaperonins', are thought to act by inhibiting folding and/or oligomerization.

During translocation, the cleavage site of the signal peptide is recognized and endoproteolytically cleaved by an integral membrane protein complex known as signal peptidase. Signal peptidase cleavage activity has been purified from dog pancreas as a complex of five distinct polypeptides with apparent molecular weights of 12, 18, 21, 22/23, and 25 kD[18]. It has also been purified from hen oviduct as a two-subunit complex of 19 and 22/23 kD[19]. Interestingly, two of the five canine signal peptidase subunits are homologous to the yeast SEC11 protein[20], thought to be a component of yeast signal peptidase since it is required for signal peptide processing and yeast cell growth[21]. These results imply that eukaryotic signal peptidase is very conserved between species. In contrast, *E. coli* signal peptidase I (leader peptidase) consists of a single subunit of 323 amino acids[22]. Although the prokaryotic and eukaryotic enzymes have a very similar substrate specificity[7], no significant identity has yet been found between *E. coli* signal peptidase I and any of the eukaryotic subunits whose sequences are known.

Recent evidence suggests that the NH_2-terminus of the signal peptide is oriented towards the cytoplasmic side of the membrane bilayer during translocation[23,24]. The active site of signal peptidase is located on the lumenal side of the membrane[25] thus requiring that at least the cleavage site domain of the preprotein must be translocated across the membrane for processing to occur. These data are consistent with the loop model put forward by Inouye[26] to explain the conformation of preproteins during translocation. This model predicts that the signal peptide forms a loop with the mature chain so that the orientation of the signal peptide is opposite that of the mature chain (which has its NH_2-terminus oriented towards the ER lumen, see Fig. 2).

What type of environment(s) are signal peptides subjected to during these processes? While it is clear that recognition by SRP and by signal peptidase involves direct protein–protein interactions, translocation across the membrane may occur through a proteinaceous pore structure. Crosslinking studies indicate that the signal peptide of preprolactin is located in the vicinity of a 34 kD integral membrane protein during its translocation[27]. This polypeptide has been termed the signal sequence receptor. However, because the *mature* domain of preprolactin also crosslinks to this 34 kD protein[28], it may serve a more general function during translocation (e.g. as a component of the pore). Nascent preprolactin can be extracted from the (ER) membrane by aqueous reagents (such as urea) or at alkaline pH[29]. This is consistent with the idea that translocation occurs through an aqueous pore. Genetic selections have identified several genes required for efficient translocation of proteins into the ER of *S. cerevisiae*[30]. The precise roles of the polypeptides encoded by these genes (e.g. SEC61, SEC62, and SEC63) have not yet been determined. However, the protein encoded by the SEC62 gene appears to be membrane associated, giving rise to speculation that it may be a component of the translocation pore and/or machinery in the ER membrane. In addition, evidence from prokaryotic systems suggests that membrane lipids may play a direct role in translocation[31]. Indeed, the hydrophobic characteristics and lack of primary sequence conservation among signal peptides are consistent with the hypothesis that interactions with lipid bilayers contribute to the regulation (mechanism) of translocation.

Preprotein Structure Required For Targeting and Translocation

A vast amount of data supports the view that signal peptides are required for translocation across the ER membrane. For example, deletion of signal peptide coding regions from genes encoding secretory proteins results in accumulation of these proteins in the cytoplasm. As discussed below, deletions and point mutations within the signal peptide can also alter translocation efficiency. While signal peptides have also been shown to be sufficient in most cases for translocation of eukaryotic preproteins, some exceptions are known[32], raising the possibility that structural compatibility between the signal peptide and its 'passenger domain' is required and/or that regions of the mature sequence aid or allow translocation. Note that signal peptides are not always present as NH_2-terminal domains which are cleaved during translocation: two members of the serine protease inhibitor family, chicken ovalbumin, and human plasminogen activator inhibitor have internal, uncleaved signal peptides.

Functional characterization of altered preproteins has provided insight about the structural requirements for targeting and translocation. Several approaches have been used: (i) genetic screens in *E. coli* or *S. cerevisiae* to identify mutations in specific preproteins which reduce their export, (ii) characterization of the functional capabilities of members of large panels of chimeric proteins generated by randomization of their prepeptide domains, and (iii) systematically altering signal sequences by site-directed mutagenesis and assaying their function *in vitro* or *in vivo*. These sequence alterations in eukaryotic signal peptides could conceivably exert their effects at one or more points along the targeting/translocation pathway including: (i) binding by SRP, (ii) release from SRP, (iii) interaction with a putative signal sequence receptor in the membrane and (iv) recognition and cleavage by signal peptidase. Rather than focusing on specific steps

in the pathway, most functional characterizations of signal peptide mutants involve analyses that do not distinguish between targeting and translocation.

Most alterations which decrease the translocation efficiency of prokaryotic and eukaryotic signal peptides tend to reduce the overall hydrophobicity or length of their h-regions (for review see ref. 6). For example, the ability of yeast preprocarboxypeptidase Y to be translocated into the secretory pathway of a mammalian cell line can be directly correlated with the hydropathic character of its h-region[33]. Deletions and point mutations which introduce charged or polar residues within the h-region often drastically reduce translocation. On the other hand, when the yeast invertase signal peptide was replaced with random amino acid sequences, about 20% were able to direct export of invertase activity while 10% directed export at essentially wild-type levels[34]. Again, a direct correlation was found between the function of these randomly generated signal peptides and their overall hydrophobicity.

A net positive charge in the signal peptide n-region has been reported to be necessary for efficient translocation across the bacterial plasma membrane. However, in eukaryotes it appears that this feature results in little, if any, functional advantage[35,36]. This may reflect: (i) differences in the mechanism of translocation between the two systems and/or (ii) the fact that bacterial preproteins have formylated amino termini while the initiator Met of eukaryotic preproteins contains a free amino group that contributes some positive charge.

Experiments have been performed to determine the requirements for signal peptide interactions with SRP. Walter and co-workers used a translational arrest assay to show that introduction of a relatively polar leucine analog, β-hydroxyleucine, into the signal peptide of preprolactin blocks its recognition by SRP[37]. It has also been shown that SRP binding is *not* dependent on the presence of a net positive charge in the n-region[35] or on any identifiable features in the c-region[38]. Taken together, these results suggest that the most critical requirement for SRP interaction is a block of hydrophobic residues in the h-region.

To identify signal peptide conformation(s) important for their function, the physical properties of isolated wild-type signal peptides have been compared to their non-functional, mutant counterparts[39]. As noted above, the wild type LamB signal peptide exhibits a strong tendency to adopt α-helix structure in environments which mimic membranes (e.g. SDS micelles and phospholipid vesicles). Some amino acid substitutions that lead to reduced LamB signal peptide function also have reduced α-helix content, whereas mutations that had little effect on function generally did not perturb α-helix formation. The substitution of charged residues in the hydrophobic core significantly reduced function. However, these charged residues did not reduce formation of α-helical structure, suggesting that α-helix formation is important but not sufficient for signal peptide function.

These results, taken together with statistical analyses of naturally occurring signal peptides, argue strongly against a system of signal peptide recognition based on primary amino acid sequence. Rather, the general physical-chemical properties of amino acids located within specific domains (e.g. hydrophobicity in the h-region) together with the ability to adopt specific secondary structures (e.g. α-helix) appear to be critical for establishing/maintaining competent function.

Preprotein Structure Required For Signal Peptidase Cleavage

The final role for a signal peptide is to act as a substrate for signal peptidase. With one known exception (bovine pregrowth hormone), signal peptidase recognizes and cleaves naturally occurring preproteins at a single site. The question of how the choice of cleavage site is made by signal peptidase, or, in effect, what governs the efficiency of cleavage, has been an enigma, given the enormous diversity in their primary structures. Because signal peptides usually contain small, neutral residues at their -3 and -1 positions, it has been suggested that these two residues form a recognition site for the enzyme[8,9]. However, because small neutral residues spaced one residue apart occur rather frequently in preproteins, other features are likely to influence the specificity of cleavage.

Our laboratory has used a model preprotein, human pre(Δpro)apo A-II, to investigate primary sequence features which regulate the site and efficiency of signal peptidase cleavage. The strength of this model is that the preprotein contains several potential sites of cleavage (e.g. after $Gly^{18}\downarrow$ and $Ala^{20}\downarrow$). Therefore, the effect of signal peptide sequence alterations can be interpreted based not only on the efficiency of cleavage but also on which cleavage site is selected. Site-saturation mutagenesis of the -1 residue of pre(Δpro)-apo A-II and expression of wild type and mutant mRNAs in a mammalian *in vitro* translation/translocation system indicated that the relative propensity for amino acids to occupy the site of cleavage is as follows: Ala>Cys>Gly>Ser, Thr>Pro>Asn, Val, Ile, Leu, Tyr, His, Arg, Asp[40]. These results agree well with the statistical distribution of these amino acids in the -1 position of naturally occurring signal peptides[8,9]. In addition to demonstrating the importance of the -1 residue in establishing a good context for signal peptidase cleavage, mutants were identified which exhibited bidirected signal peptidase cleavage.

The possibility that other features besides the -1 residue influence the site and efficiency of signal peptidase cleavage was investigated using pre(Δpro)-apo A-II and its derived cleavage site (-1 residue) mutants. The position of the signal peptide h/c boundary was systematically altered to determine if signal peptidase tended to conserve a given distance

from the COOH-terminus of the hydrophobic core to the cleavage site[41]. For mutants which exhibited bidirected cleavage after $Gly^{18\downarrow}$ and $Xaa^{20\downarrow}$, movement of the predicted position of the h/c boundary downstream (i.e. COOH-terminal towards $Xaa^{20\downarrow}$) caused more cleavage to occur at position 20 while more cleavage shifted to position 18 when the h/c boundary was moved upstream (NH_2-terminal; away from the $Xaa^{20\downarrow}$ position). The data indicated that signal peptidase prefers a distance of 4–5 residues from the h/c boundary to the site of cleavage. However, in some cases the physical-chemical characteristics of potential −1 residues COOH-terminal to the hydrophobic core are capable of overriding the tendency to conserve this 'optimal' distance. Interestingly, when the length of the h-region of pre(Δpro)apo A-II was expanded from its NH_2-terminal end, the position of signal peptidase cleavage shifted to the NH_2-terminal position[56]. This result suggests that the position of the n/h boundary can also influence the site of signal peptidase cleavage (located 16 amino acids away). Related observations have been made from study of the polar, cytoplasmic domains NH_2-terminal to the uncleaved signal-membrane anchor sequences of two human type II membrane proteins – invariant chain (Iγ)[42] and asialoglycoprotein receptor H1[43]. When these cytoplasmic domains were deleted from cDNAs expressed in mammalian cells, signal peptidase cleavage occurred at the COOH-terminal end of the previously uncleaved signal-anchor sequences. The manner in which signal peptidase cleavage is regulated by the NH_2-terminal region of signal peptides is not known but alterations in this region conceivably could change the topology (positioning) of the signal peptide in the membrane (thereby affecting presentation of the cleavage site) and/or alter direct protein–protein interactions between the peptidase and the n-region. Finally, alteration of sequences located COOH-terminal to signal peptidase cleavage sites has been shown to influence the site[44,48] and efficiency[45] of signal peptidase cleavage.

What is the relationship between translocation and signal peptidase cleavage? Sequence alterations in the vicinity of the cleavage site can block cleavage while allowing translocation to occur[38,46]. Such 'uncoupling' of co-translational translocation from cleavage suggests that proteolytic processing is not a prerequisite for translocation and that structural features in signal peptides required for both functions can be distinguished functionally. However, Cioffi *et al.* recently found that dipeptide deletions across the h-region of the preproparathyroid hormone signal peptide had parallel effects on co-translational translocation/signal peptidase processing and post-translational cleavage by purified hen oviduct signal peptidase; features located at the NH_2-terminus of the hydrophobic core reduced translocation and processing while deletions at the COOH-terminal region of the core had no deleterious effects[47]. Based on these results, a two subunit model for signal peptidase cleavage was proposed in which a non-catalytic subunit binds the hydrophobic core and a catalytic subunit interacts with the cleavage site[47]. Such an interaction between signal peptidase and the hydrophobic core would be consistent with our observation that NH_2-terminal expansion of the h-region of pre(Δpro)apo A-II results in a shifting of cleavage to a more NH_2-terminal position. We have found that the extent of hydrophobicity in the vicinity of the h/c boundary of pre(Δpro)apo A-II directly correlates with the extent of *in vitro* co-translational translocation[48]. We also find such a correlation when these mutants are cleaved *post*-translationally by purified hen oviduct signal peptidase. These results suggest that there is an 'overlap' at the h/c boundary between features required for efficient translocation and cleavage. The fact that signal peptidase is a tight complex of multiple polypeptides and that structural features in prepeptides required for both cleavage and translocation exist, raise the possibility that one or more of the polypeptides which co-purify with signal peptidase activity may be involved in the process of translocation.

Future Directions

Analysis of the relationships between signal peptide structure and function has been the focus of considerable research effort in the past few years. More knowledge about how these protein domains 'work' would benefit those who wish to ensure proper processing and secretion of genetically engineered proteins expressed in recombinant systems. This information may have possible implications for understanding other biological processes which involve cellular recognition of protein domains that vary in primary sequence, e.g. the binding of peptide antigens by the major histocompatibility complex and the binding of newly translocated proteins in the ER by heavy chain binding protein (BiP).

In the future it will be necessary to further characterize the interactions of signal peptides with cellular components involved in co-translational translocation and processing. More detailed characterization of the translocation machinery will be required to fully understand signal peptide structure/function relationships as well as how the mature polypeptide domain passes through the membrane. Information about the interactions of signal peptides with these components (e.g. SRP) should provide additional model systems for evaluating how signal peptidase can accurately cleave its many diverse substrates. For example, it is not known whether signal peptidase specificity is influenced by signal peptide topology in the membrane or is due solely to enzyme–substrate interactions. Comparative analysis of the sites of (i) post-translational cleavage of model preproteins (containing systematic alterations) by purified signal peptidase and (ii) co-translational cleavage by intact, translocation-component microsomal membrane vesicles with associated signal peptidase

activity should help address this issue. Moreover, the availability of purified signal peptidase will allow kinetic and biochemical studies of enzyme–substrate interactions to be performed. These should also shed additional light on the question of its substrate specificity.

References

1 Blobel, G. (1980). Intracellular protein topogenesis. *Proc. Natl Acad. Sci. USA* **77**, 1496–1500.

2 Wickner, W. and Lodish, H. F. (1985). Multiple mechanisms of protein insertion into and across membranes. *Science* **230**, 400–407.

3 Milstein, C., Brownlee, G. G., Harrison, T. M. and Mathews, M. B. (1972). A possible precursor of immunoglobin light chains. *Nature* **239**, 117–120.

4 Blobel, G. and Dobberstein, B. (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J. Cell Biol.* **67**, 835–851.

5 Blobel, G. and Dobberstein, B. (1975). Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *J. Cell Biol.* **67**, 852–862.

6 Gierasch, L. M. (1989). Signal sequences. *Biochemistry* **28**, 923–930.

7 Watts, C., Wickner, W. and Zimmerman, R. (1983). M13 procoat and pre-immunoglobulin share processing specificity but use different membrane receptor mechanisms. *Proc. Natl Acad. Sci. USA* **80**, 2809–2813.

8 von Heijne, G. (1985). Signal sequences: the limits of variation. *J. Mol. Biol.* **184**, 99–105.

9 Perlman, D. and Halvorson, H. O. (1983). A putative signal peptidase recognition site and sequences in eukaryotic and prokaryotic signal peptides. *J. Mol. Biol.* **167**, 391–409.

10 Reddy, G. L. and Nagaraj, R. (1989). Circular dichroism studies on synthetic signal peptides indicate β-conformation as a common structural feature in highly hydrophobic environment. *J. Biol. Chem.* **264**, 16591–16597.

11 Bruch, M. D., McKnight, C. J. and Gierasch, L. M. (1989). Helix formation and stability in a signal sequence. *Biochemistry* **28**, 8554–8561.

12 Walter, P. and Lingappa, V. R. (1986). Mechanism of protein translocation across the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **2**, 499–516.

13 Bernstein, H. D., Poritz, M. A., Strub, K., Hoben, P. J., Brenner, S. and Walter, P. (1989). Model for signal sequence recognition from amino acid sequence of 54K subunit of signal recognition particle. *Nature* **340**, 478–486.

14 Connolly, T. and Gilmore, R. (1989). The signal recognition particle receptor mediates the GTP-dependent displacement of SRP from the signal sequence of the nascent polypeptide. *Cell* **57**, 599–610.

15 Rothman, J. E. (1989). GTP and methionine bristles. *Nature* **340**, 433–434.

16 Jurnak, F., Heffron, S. and Berghann, E. (1990). Conformational changes involved in the activation of ras p21: implications for related proteins. *Cell* **60**, 525–528.

17 Meyer, D. I. (1988). Preprotein conformation: the year's major theme in translocation studies. *Trends Biochem. Sci.* **13**, 471–474.

18 Evans, E. A., Gilmore, R. and Blobel, G. (1986). Purification of microsomal signal peptidase as a complex. *Proc. Natl Acad. Sci.* **83**, 581–585.

19 Baker, R. K. and Lively, M. O. (1987). Purification and characterization of hen oviduct microsomal signal peptidase. *Biochemistry* **26**, 8561–8567.

20 Shelness, G. S. and Blobel, G. (1990). Two subunits of the canine signal peptidase complex are homologous to yeast SEC11 protein. *J. Biol. Chem.*, in press.

21 Böhni, P. C., Deshaies, R. J. and Schekman, R. W. (1988). SEC11 is required for signal peptide processing and yeast cell growth. *J. Cell Biol.* **106**, 1035–1042.

22 Wolfe, P. B., Wickner, W. and Goodman, J. M. (1983). Sequence of the leader peptidase gene of *Escherichia coli* and the orientation of leader peptidase in the bacterial envelope. *J. Biol. Chem.* **258**, 12073–12080.

23 Kuhn, A. (1987). Bacteriophage M13 procoat protein inserts into the plasma membrane as a loop structure. *Science* **238**, 1413–1415.

24 Shaw, A. S., Rottier, P. J. M. and Rose, J. K. (1988). Evidence for the loop model of signal-sequence insertion into the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA* **85**, 7992–7996.

25 Jackson, R. C. and Blobel, G. (1977). Post-translational cleavage of presecretory proteins with an extract of rough microsomes from dog pancreas containing signal peptidase activity. *Proc. Natl Acad. Sci. USA* **74**, 5598–5602.

26 Inouye, M. and Halegoua, S. (1980). Secretion and membrane localization of proteins in *Escherichia coli*. *CRC Crit. Rev. Biochem.* **7**, 339–371.

27 Weidmann, M., Kurzchalia, E., Hartmann, E. and Rapoport, T. A. (1987). A signal sequence receptor in the endoplasmic reticulum membrane. *Nature* **328**, 830–833.

28 Weidmann, M., Goerlich, D., Hartmann, E., Kurzchalia, T. V. and Rapoport, T. A. (1989). Photocrosslinking demonstrates proximity of a 34 kDa membrane protein to different portions of preprolactin during translocation through the endoplasmic reticulum. *FEBS Lett.* **257**, 263–268.

29 Gilmore, R. and Blobel, G. (1985). Translocation of secretory proteins across the microsomal membrane occurs through an environment accessible to aqueous perturbants. *Cell* **42**, 497–505.

30 Rothblatt, J. A., Deshaies, R. J., Sanders, S. L., Daum, G. and Schekman, R. (1989). Multiple genes are required for proper insertion of secretory proteins into the endoplasmic reticulum in yeast. *J. Cell Biol.* **109**, 2641–2652.

31 de Vrije, T., de Swart, R. L., Dowhan, W., Tommassen, J. and de Kruijff, B. (1988). Phosphatidylglycerol is involved in protein translocation across *Escherichia coli* inner membranes. *Nature* **334**, 173–175.

32 Moreno, F., Fowler, A. V., Hall, M., Silhavy, T. J., Zabin, I. and Schwartz, M. (1980). A signal sequence is not sufficient to lead β-galactosidase out of the cytoplasm. *Nature* **286**, 356–359.

33 Bird, P., Gething, M.-J. and Sambrook, J. (1990). The functional efficiency of a mammalian signal peptide is directly related to its hydrophobicity. *J. Biol. Chem.* **265**, 8420–8425.

34 Kabsch, C. A., Preuss, D., Grisafi, P. and Botstein, D. (1987). Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science* **235**, 312–317.

35 Garcia, P. D., Gravater, J., Inouye, M. and Walter, P. (1987). Wild-type and mutant signal peptides of *Escherichia coli* outer membrane lipoprotein interact with equal efficiency with mammalian signal recognition particle. *J. Biol. Chem.* **262**, 9463–9468.

36 Nothwehr, S. F. and Gordon, J. I. (1990). Structural features in the NH_2-terminal region of a model eukaryotic signal peptide influence the site of its cleavage by signal peptidase. *J. Biol. Chem.*, in press.

37 Walter, P., Ibrahimi, I. and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum membrane. I. Signal recognition protein (SRP) binds to *in vitro* assembled polysomes synthesizing secretory protein. *J. Cell Biol.* **91**, 545–550.

38 Nothwehr, S. F., Folz, R. J. and Gordon, J. I. (1989). Uncoupling of co-translational translocation from signal peptidase processing in a mutant rat preapolipoprotein A-IV with a deletion that includes the COOH-terminal region of its signal peptide. *J. Biol. Chem.* **264**, 4642–4647.

39 McKnight, C. J., Briggs, M. S. and Gierasch, L. M. (1989). Functional and nonfunctional LamB signal sequences can be distinguished by their biophysical properties. *J. Biol. Chem.* **264**, 17293–17297.

40 Folz, R. J., Nothwehr, S. F. and Gordon, J. I. (1988). Substrate specificity of signal peptidase: site-saturation mutagenesis at position −1 regulates cleavage between multiple sites in human pre(Δpro)apo A-II. *J. Biol. Chem.* **263**, 2070–2078.

41 Nothwehr, S. F. and Gordon, J. I. (1989). Eukaryotic signal peptide structure/function relationships: identification of conformational features which influence the site and efficiency of co-translational proteolytic processing by site-directed mutagenesis of human pre(Δpro)apo A-II. *J. Biol. Chem.* **264**, 3979–3978.

42 Lipp, J. and Dobberstein, B. (1986). The membrane spanning segment of invariant chain (Ii) contains a potentially cleavable signal sequence. *Cell* **46**, 1103–1112.

43 Schmid, S. R. and Spiess, M. (1988). Deletion of the amino-terminal domain of asialoglycoprotein receptor H1 allows cleavage of the internal signal sequence. *J. Biol. Chem.* **263**, 16886–16892.

44 Folz, R. J. and Gordon, J. I. (1986). Deletion of the propeptide from human preproapolipoprotein A-II redirects co-translational processing by signal peptidase. *J. Biol. Chem.* **261**, 14752–14759.

45 Andrews, D. W., Perara, E., Lesser, C. and Lingappa, V. R. (1988). Sequences beyond the cleavage site influence signal peptide function. *J. Biol. Chem.* **263**, 15791–15798.

46 Hortin, G. and Boime, I. (1981). Transport of an uncleaved preprotein into the endoplasmic reticulum of rat pituitary cells. *J. Biol. Chem.* **256**, 1491–1494.

47 Cioffi, J. A., Allen, K. L., Lively, M. O. and Kemper, B. (1989). Parallel effects of signal peptide hydrophobic core modifications on co-translational and post-translational cleavage by purified signal peptidase. *J. Biol. Chem.* **264**, 15052–15058.

48 Nothwehr, S. F., Hoeltzli, S. D., Allen, K. L., Lively, M. O. and Gordon, J. I. (1990). Residues flanking the COOH-terminal c-region of a model eukaryotic signal peptide influence the site of its cleavage by signal peptidase and the extent of coupling of its co-translational translocation and proteolytic processing *in vitro*. *J. Biol. Chem.*, in press.

Steven F. Nothwehr and **Jeffrey I. Gordon*** are at the Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St Louis, Missouri 63110, USA.
* To whom correspondence should be sent.